Identification and characterisation of inner membrane components of TolC-dependent secretion mechanisms in the cyanobacterium *Synechocystis* sp. PCC 6803

Cátia Filipa Correia Gonçalves
Mestrado em Biologia Celular e Molecular
Departamento de Biologia
2016

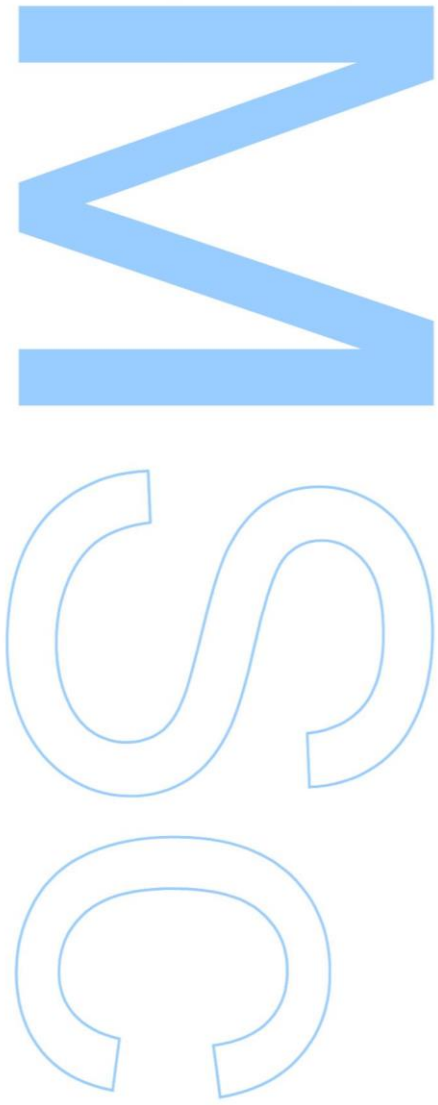
Orientador
Professora Doutora Paula Tamagnini, Professora Associada da Faculdade de Ciências da Universidade do Porto, Líder de Grupo no i3S/IBMC, Universidade do Porto

Coorientador
Doutor Paulo Oliveira, Bolseiro de Pós-Doutoramento no i3S/IBMC, Universidade do Porto

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____



O caos é uma ordem por decifrar.

José Saramago

Agradecimentos

Agradeço à Professora Paula Tamagnini por todas as oportunidades que me proporcionou ao longo deste ano e pela disponibilidade apresentada durante o processo de escrita e correção desta dissertação.

Paulo, agradeço-te por tudo o que me ensinaste, pela paciência e pela boa disposição com que sempre me prendaste. Acima de tudo, agradeço-te pelo apoio nos momentos mais difíceis, por me teres ensinado a (tentar) ver o lado positivo das coisas, mesmo quando tudo corre mal. Agradeço-te ainda por teres lido esta dissertação mais vezes do que aquilo que é recomendável e por nunca deixares de me educar com os teus comentários, sempre construtivos. O meu muito obrigada.

Aos membros do grupo “Bioengineering and Synthetic Biology” agradeço todos os bons momentos de trabalho, a partilha de ideias e, claro está, os momentos de galhofa, quer seja ao almoço ou ao lanche. Não irei fazer menções especiais, pois aprendi algo com cada um de vós e nunca vos faltou o entusiasmo para me ajudarem sempre que precisei. Sem a vossa companhia, esta jornada teria sido muito mais difícil.

Agradeço ao grupo “Parasite Disease” (i3S/IBMC), liderado pela Professora Doutora Anabela Cordeiro da Silva, a disponibilização do equipamento Synergy. Foi um contributo fundamental para o sucesso deste trabalho.

Inês Proença, sei que faz já algum tempo que as circunstâncias nos voltaram a separar, mas não poderia deixar de te agradecer por teres sido a melhor companheira que alguém poderia pedir durante aqueles penosos anos de licenciatura. Guardo boas recordações, o tempo que partilhamos foi precioso. Saudades!

Sofia Oliveira, a minha comunicadora de Ciência preferida, a tua companhia, a tua boa disposição, as tuas histórias e todos os teus comentários às minhas apresentações (das quais raramente gostas) foram essenciais ao longo destes anos. Aguardo com ânsia a gentil oferta de um autógrafo antes que as luzes da ribalta te arrebatem.

Família, agradeço-vos o apoio incondicional, as palavras de apreço e todos os gestos de ternura por mais insignificantes que possam ter parecido. Nunca me falharam quando mais precisei e poucas são as palavras que possam expressar o quão grata estou. Sem dúvida alguma que esta jornada teria sido praticamente impossível sem o vosso contributo.

E por último, agradeço aos meus pais pelo amor incondicional, por me terem concedido esta e tantas outras oportunidades, por nunca duvidarem das minhas capacidades e por nunca me terem deixado desistir. Foi convosco que aprendi o significado da palavra perseverança. Obrigada por terem estado sempre presentes, por me ouvirem mesmo quando o tema é demasiadamente estranho para ser entendido. Nunca teria chegado até aqui sem o vosso amor. Amo-vos!

Abstract

Synechocystis sp. PCC 6803 is one of the most widely studied cyanobacteria, presenting minimal growth requirements, a relatively simple genetic background and a fairly complete toolbox for genetic manipulation. This unicellular cyanobacterium has a multi-layered cell envelope consisting of an inner membrane (IM), a peptidoglycan layer, and an outer membrane, which defines the cell boundaries and controls the entry and exit of numerous compounds. Consequently, *Synechocystis* has evolved multiprotein complexes capable of crossing this physical barrier. Multiprotein transport systems have been thoroughly studied in proteobacteria. However, knowledge about these assemblies in cyanobacteria remains scarce, even regarding the identity of the molecular components of IM translocase complexes. In the present work, a bioinformatics analysis was carried out aiming at identifying IM translocase partners of TolC-dependent export mechanisms. The 11 candidate genes identified were inactivated by double homologous-recombination, and their function was studied by phenotypic analysis. To investigate the role of the proteins encoded by the candidate genes in multidrug efflux mechanisms, *Synechocystis* mutants were exposed to compounds with antimicrobial activity – chloramphenicol, sodium dodecyl sulphate and Triton X-100. Moreover, the efflux activity of the wild-type and mutant strains was also evaluated upon exposure to the fluorophore ethidium bromide. Finally, the involvement of the candidate proteins in protein secretion was studied by analysing the extracellular protein content present in the growth media, which was further complemented by observing negatively stained cells by transmission electron microscopy. The combination of computational screening with the experimental work described, provided sufficient evidence to identify components involved in multidrug efflux and protein secretion. Remarkably, the results obtained suggest the participation of five proteins in a TolC-dependent protein secretion system, which had not yet been reported in any other Gram-negative bacteria. In addition, knockout mutants lacking functional subunits of TolC-dependent multidrug efflux pumps showed significant modifications in the post-translational processing of pilins. To the best of our knowledge, this is the first time that such IM translocase complexes are implicated with pili maturation and structure. In conclusion, this work represents the first approach to the study of TolC-dependent export mechanisms in *Synechocystis*, and highlights the intricate mechanisms of function overlap between IM components, and a TolC-independent functional role among them.

Keywords: *Synechocystis*, TolC, RND, T1SS, Dev

Resumo

Synechocystis sp. PCC 6803 é uma das cianobactérias mais estudadas, apresentando requisitos nutricionais mínimos, um genoma relativamente simples e um conjunto razoável de ferramentas que possibilitam a sua manipulação genética. Esta cianobactéria possui um invólucro celular composto por três estruturas distintas, uma membrana interna (MI), uma camada de peptidoglicano e uma membrana externa, que definem os limites da célula e controlam a entrada e saída de inúmeros compostos. Consequentemente, *Synechocystis* desenvolveu complexos transportadores multiproteicos que são capazes de transportar substratos através desta barreira física. Embora estes sistemas tenham vindo a ser meticulosamente estudados em proteobactérias, o conhecimento acerca destes em cianobactérias é escasso, nomeadamente no que diz respeito aos componentes dos complexos translocase da MI. Neste trabalho, foi realizada uma análise bioinformática com o objetivo de identificar subunidades dos complexos translocase relacionados com mecanismos de transporte dependentes do TolC. Os 11 genes candidatos identificados foram inativados através de recombinação homóloga dupla e a função das proteínas por eles codificadas foi estudada através de análise fenotípica. Para investigar o papel destas proteínas em mecanismos de secreção, os mutantes foram expostos a compostos com atividade antimicrobiana – cloramfenicol, dodecil sulfato de sódio e Triton X-100. Para além disso, a atividade de efluxo da diferentes estirpes foi avaliada após exposição ao fluorocromo brometo de etídio. Por fim, o envolvimento destas proteínas em mecanismos de secreção de proteínas foi estudado através da análise do conteúdo proteico extracellular e por observações da superfície celular no microscópio eletrónico de transmissão. A combinação de técnicas bioinformáticas com o trabalho experimental desenvolvido, providenciou resultados suficientes para a identificação de componentes envolvidos em mecanismos de secreção de compostos e proteínas. Os resultados obtidos sugerem a participação de cinco proteínas num sistema de secreção dependente do TolC, o que ainda não havia sido reportado em outras bactérias Gram-negativas. Para além disso, mutantes deficientes em proteínas relacionadas com mecanismos de secreção de compostos apresentam modificações significativas no processamento pós-tradução das pilinas, sendo que esta é a primeira vez que tais complexos são implicados na maturação dos pili. Em conclusão, os resultados aqui apresentados representam uma primeira abordagem ao estudo dos sistemas de transporte dependentes do TolC em *Synechocystis* e realçam a existência de uma sobreposição de função em alguns componentes, bem como funções independentes do TolC.

Palavras-chave: *Synechocystis*, TolC, RND, T1SS, Dev

Table of contents

List of figures

List of tables

List of abbreviations

ABC	ATP-binding cassette
ATP	Adenosine triphosphate
BLAST	Basic Alignment Search Tool
CDD	Conserved Domain Database
CFU	Colony forming unit
C-terminal	Carboxyl terminal
Dev	Heterocyst development
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
Gsp	General secretion pathway
IC_{50}	Half maximal growth inhibitory concentration
HAE1	Hydrophobe/amphiphile efflux-1
HME	Heavy metal efflux
IM	Inner membrane
IMC	Inner membrane component
MATE	Multidrug and toxic compound extrusion
MDR	Multidrug efflux
MFP	Membrane fusion protein
MFS	Major facilitator superfamily
MIC	Minimum inhibitory concentration
MMPL	Mycobacterial membrane protein large
MP-domain	Membrane proximal domain
NBD	Nucleotide-binding domain
N-terminal	Amino-terminal
OD	Optical density
OEP	Outer membrane efflux protein

OM	Outer membrane
OMP	Outer membrane protein
OMV	Outer membrane vesicles
ORF	Open reading frame
PAP	Periplasmic adaptor protein
PAS	Periodic acid-Schiff
PCR	Polymerase chain reaction
PMF	Proton motive force
RND	Resistance-nodulation-division
ROS	Reactive oxygen species
RTX	Repeats-in-toxin
SDS	Sodium dodecyl sulphate
S-layer	Surface layer
SMR	Small multidrug resistance
SPaseII	Signal peptidase II
T1SS	Type I secretion system
T2SS	Type II secretion system
T4SS	Type IV secretion system
TEM	Transmission electron microscopy

1. INTRODUCTION

1 Introduction

1.1 Cyanobacteria: an ancient group of prokaryotes

Cyanobacteria are one of the oldest and morphologically most diverse phylum on our planet. These prokaryotic organisms have had a pivotal role in the history of life on Earth, being the first organisms to ever develop the ability to perform oxygenic photosynthesis (*1*). In addition, they are considered to be responsible for the sharp rise of atmospheric oxygen that provided the basis for the evolution of aerobic respiration and more complex life forms, forever reshaping life on the planet (*1*, *2*). Some cyanobacteria are also capable of performing the biological reduction of atmospheric nitrogen (N_2), actively contributing to the carbon and nitrogen cycles at the ecosystem level (*2*, *3*). Over their long evolutionary history, cyanobacteria have developed extensive morphological diversity, representing one of the most heterogeneous group of prokaryotes (*2*, *4*). Their great adaptability and plasticity allowed cyanobacteria to conquer and thrive in numerous habitats (e.g. marine, freshwater and terrestrial environments) and even endure extreme abiotic stresses (*3*). Moreover, the metabolic diversity of cyanobacteria mirrors the diversity of habitats, and these photoautotrophic microorganisms are seen as a prolific source of biologically active compounds with antiviral, antibacterial, antifungal and anticancer activities recognised for many decades (*5*).

Ever since it was demonstrated the natural ability of some cyanobacteria to incorporate exogenous DNA (*6*, *7*), these photoautotrophic microorganisms have been employed in the study of cell division and differentiation mechanisms, photosynthesis, N_2 fixation, hydrogen metabolism, and in a number of other areas of basic and applied research. In the last couple of decades, progresses in metabolic engineering, and synthetic and systems biology, have allowed the development of engineered cyanobacteria for the production of numerous commercially relevant chemicals compounds, from short chains alcohols (e.g. ethanol) (*8*, *9*) to plant secondary metabolites (e.g. isoprene and ethylene) (*10*) and hydrocarbons (e.g. limonene) (*11*). However, cyanobacterial cell factories still present low production titers when compared to other microorganisms (*12*). For example, photoautotrophic production of ethanol in *Synechocystis* sp. PCC 6803 is seven times lower than the production of ethanol from glucose in *Escherichia coli* (*8*, *13*). Even though major progresses have been made over the past decades, only a few compounds produced by cyanobacteria have reached a commercial scale, namely pigments (phycobiliproteins and carotenoids), polysaccharides (Nostoflan), vitamins (vitamin B and E) and enzymes (amylases and proteases) (*14*, *15*).

Up to the present time, research efforts have been mainly centralised on the design and optimisation of synthetic metabolic pathways and their establishment in the host organism. However, the challenges presented by these bioprocesses imply that several strategies must be carefully considered to improve production titers and overcome the major difficulties on the path to commercialisation (*16*, *17*). To maximise the production of a particular metabolite it is required a proper understanding of how living cells control the rate flux of substrates through metabolic pathways (*18*). Cellular transporters contribute significantly to the control of metabolic fluxes for they mediate the influx of substrates and efflux of products or potentially cytotoxic compounds (*19*). These steps are often disregarded in the field of biotechnology, despite being considerably rate-controlling (*19*). In most cases, intracellular accumulation of the target compound or metabolic by-products might influence negative feedback loops or inhibit growth, hindering the production of the compound of interest (*16*, *20*). Moreover, to simplify recovery procedures it is advantageous that the target chemicals produced in the cytoplasm are excreted into the extracellular medium through adequate transporters (*20*). On the one hand, this perception has led to the discovery of many efflux transporters for relevant chemical compounds, such as hydrocarbons (*21*) and fatty acids (*22*). On the other hand, directed evolution approaches have been successfully applied to explore the potential of broad-substrate export machineries. For example, an engineered plasma membrane transporter from *E. coli* has improved tolerance to the non-native substrate n-butanol, enhancing growth rates in the presence of this biofuel up to 25% (*23*). Engineering the same transporter has also improved efflux efficiency for n-octane up to 51% in *E. coli* (*24*). Therefore, optimised expression of efflux transporters can lead to increased production yields, improved product tolerance of the biocatalyst cell, and maximised product recovery from the extracellular space (*16*, *20*).

1.2 The Gram-negative cell envelope

Even though the origin of life and the first cell remain a matter of speculation, it is accepted that prebiotic systems capable of replicating molecules existed long before the first cell. In fact, it is the inception of membranes that marks the boundary between life and nonlife (*25*). The first cell had a clear advantage over prebiotic systems, as chemical reactions could occur more efficiently in an enclosed space and the membrane could also serve as a selective barrier that controlled the entry and exit of compounds; thus, membranes were an important evolutionary step towards optimised cellular metabolism (*26*). All present-day cells are surrounded by membranes, and bacterial cells, which often have to face unpredictable and hostile environments, have evolved a complex and sophisticated cell envelope that protects them (*27*). Gram-negative bacteria have a

multi-layered cell envelope consisting of an inner membrane (IM), a peptidoglycan layer and an outer membrane (OM), as schematically represented in **Figure 1**.

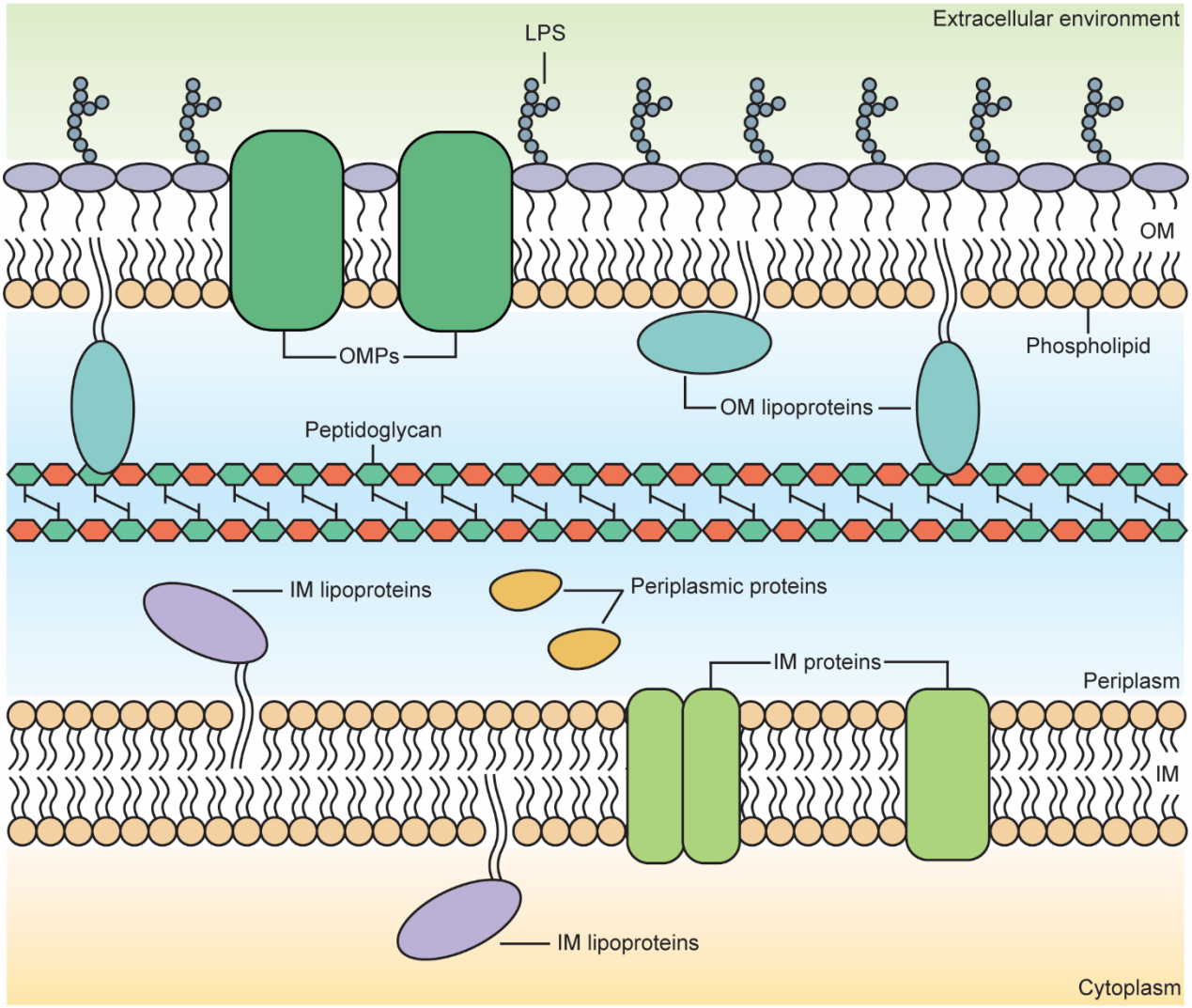


Figure 1: Structure of the Gram-negative cell envelope.

In Gram-negative bacteria, the cytoplasm is surrounded by the inner membrane (IM) or plasma membrane, a phospholipid bilayer that also contains proteins. The periplasm is the aqueous compartment enclosed by the IM and the outer membrane (OM); it contains soluble proteins and the peptidoglycan (PG) layer. The OM is anchored to the rest of the cell through proteins that are covalently attached to the PG. The OM is asymmetric as it contains phospholipids in the inner leaflet and lipopolysaccharides (LPS) in the outer leaflet. Additionally, the OM also contains two types of proteins, integral outer membrane proteins (OMP) and lipoproteins. Adapted from Silhavy et *al.*, 2010 (*27*).

Many bacteria also have an outermost layer, the surface layer (S-layer), which is composed by a single protein that surrounds the cell (*27*). Nonetheless, it should be noted that cell envelopes may differ in species and environmentally specific ways. Cyanobacteria are classified as Gram-

negative microorganisms, despite the presence of some features typical of Gram-positive bacteria in their cell envelope, namely: (i) the cyanobacterial peptidoglycan layer is considerably thicker (10-700 nm) than that of most Gram-negative bacteria, exhibits a higher degree of cross-linking and can be further complexed by the presence of specific polysaccharides, (ii) the OM of cyanobacteria contains constituents that are not usually present in Gram-negative bacteria, such as carotenoids, unusual fatty acids, or porins anchored to the peptidoglycan layer (*28*, *29*).

1.3 Transport systems in Gram-negative bacteria

As a result of the multi-layered structure of the cell envelope, import and export mechanisms are more intricate in Gram-negative bacteria than in Gram-positive microorganisms. Active transport processes occur across the IM and the energy is provided by cellular adenosine triphosphate (ATP) or by ion gradients across the membrane (*30*). In contrast, transport across the OM occurs mainly by diffusion, following a concentration gradient between the periplasmic space and the extracellular environment, as active transport across the OM is only possible by interaction with energised proteins from the IM (*31*). Furthermore, it has been estimated that the periplasmic space is too wide (at least 130 Å) to allow direct interaction between most OM and IM proteins (*32*). Consequently, sophisticated and complex transport machineries are required to overcome this mechanical and physical barrier and define a course through the different layers; therefore, Gram-negative bacteria have evolved multiprotein complexes that span the width of the cell envelope, constituting an effective transport system (*31*). The simplest multiprotein assembly comprises an IM component (IMC), a periplasmic adaptor protein (PAP), previously known as membrane fusion protein (MFP), and an OM protein (OMP) channel, which belongs to the outer membrane efflux protein (OEP) family (*31*, *33*). This double-membrane-spanning assembly uses a one-step export mechanism, whereby substrates are directly translocated from the bacterial cytoplasm to the extracellular environment (*34*, *35*). Alternatively, structural studies indicate that in some cases the transporter is also capable of capturing the substrates directly from the periplasm or from the outer leaflet of the plasma membrane (*36*).

1.3.1 Multiprotein complexes: transport through the TolC channel

TolC (tolerance to colicin E1) is the archetypal member of the OEP family that includes proteins central to two distinct export systems: the type I protein secretion system (T1SS) and the multidrug efflux (MDR) pumps (*31*). TolC homologues are found in almost all Gram-negative bacteria, which confirms the importance of this OMP for cells (*32*). Some bacteria contain multiple copies of OEPs in their genome (e.g. *Pseudomonas*), whereas others encode a single TolC

channel (e.g. enterobacteria), which couples with multiple IM transporters, carrying out a broad range of physiological purposes (*37*). Phylogenetic analyses of TolC homologues, have shown that OEPs can be grouped according to their export function: (i) multidrug efflux, (ii) cation efflux, and (iii) protein secretion (*32*). Therefore, OEPs are involved in the translocation of a wide variety of compounds: from complex endogenous molecules (proteins, glycolipids or fatty acids) to exogenous compounds that reach the cytoplasm (antibiotics, detergents, dyes, or other cytotoxic molecules) (*31*, *32*).

OEP sequences are highly divergent and have in common only two conserved domains, which are tandemly repeated (*33*). In spite of the low sequence homology, the crystal structures of the several OEPs solved thus far underpin a remarkable structural similarity, and have provided many notable insights into the mechanistic and function of these homotrimeric protein channels (*38*, *39*). OEPs assemble as an elongated homotrimer and have three distinct structural domains: a transmembrane β-barrel, a periplasmic α-helical barrel, and mixed α/β-fold, known as equatorial domain, also in the periplasm (*38*). In TolC, the distal (upper) end of the structure is open (20 Å in diameter) while the proximal (lower) end narrows to a virtual close (2 Å in diameter), due to the pairs of helices arranged into coiled coils (*39*). This arrangement of coiled-coils hairpins arises from the fact that each protomer has a structural repeat (*33*). Even though the overall architecture is conserved across different TolC homologues, some OEPs, namely OprM, also present a flexible N-terminal tail, which is often lipidated and inserted in the OM (*38*).

Transport through the TolC channel-tunnel is achieved in a single, energy-coupled step. In this mechanism, a continuous channel is created through the assembly of a multiprotein complex that spans the periplasmic space and both the IM and OM (*32*, *39*). The underlying assembly process is determined by the direct interaction of an IM translocase of two proteins (IMC and PAP), which provides substrate specificity and energy, with the third located in the OM – a protein belonging to the OEP family, in this case TolC (*38*, *39*). The tripartite complex works as a transient translocation system, wherein the IM translocase bound to its substrate recruits the OMP to form the active export structure, and as soon as the substrate is exported, the machinery is disengaged and the components revert to their initial state (*39*). TolC and its homologues are capable of interacting with different translocase complexes, whereas the IMC and the PAP usually recognise a specific set of substrates (*31*, *32*). Promiscuous partnering confers an advantage for it allows TolC to export a large variety of molecules from the cytoplasm (*32*). Furthermore, it has been shown that tripartite efflux systems often display a cooperative interaction between them and can act sequentially when one fails, which greatly broadens the spectrum of possibilities (*36*).

1.3.1.1 TolC-dependent multidrug efflux

In the course of their evolution, bacteria have established different strategies to withstand exposure to antimicrobial compounds. MDR pumps are one of the most frequently employed resistance strategies because they help to reduce the intracellular drug concentration to subtoxic levels, playing an important role in the innate resistance of Gram-negative bacteria (*31*). However, their physiological role stretches well beyond detoxification of cytotoxic molecules and antibiotic resistance, as MDR pumps are also relevant for bacterial virulence, cell homeostasis and intercellular communication (*40*, *41*). There are at least five families of IM multidrug transporters that are associated to MDR efflux pumps (**Figure 2**): the ATP-binding cassette (ABC) superfamily, the major facilitator superfamily (MFS), the small multidrug resistance superfamily (SMR), the resistance-nodulation-division (RND) superfamily and the multidrug and toxic compound extrusion (MATE) superfamily, which rely on different energy sources to catalyse transport (*31*). ABC transporters use ATP as an energy source to drive transport across the IM, whereas the other four families rely on an electrochemical gradient (*35*). Specifically, MFS, SMR and RND proteins employ the proton (H^+) motive force (PMF), and the MATE superfamily is characterised by either sodium (Na^+) or H^+ antiport system (*31*, *35*).

TolC interacts with proteins belonging to the RND, ABC and MFS superfamilies (*36*, *38*). The RND transporter superfamily was the first group of transporters associated with TolC for which crystal structures became available, influencing early models of assembly and providing insightful glimpses into these multiprotein complexes. RND pumps are specific for toxic molecules and can be divided into two different subfamilies according to their substrates, the multidrug hydrophobe/amphiphile efflux-1 (HAE1) family and the heavy metal efflux (HME) family (*36*). Typically, HAE1 transporters are trimeric assemblies, with each protomer consisting of a transmembrane domain containing twelve transmembrane α-helices and two large hydrophilic loops that comprise the substrate-binding porter domain and the OMP-coupling docking domain. The HME pumps have a very similar trimeric assembly, while the general protomer architecture is also shared with the SecDF family involved in protein secretion (*36*, *38*). OEPs and IM transporters have been thoroughly studied. Nevertheless, the role of PAPs in these tripartite assemblies has remained elusive as the crystal structures of several adaptor proteins were only recently solved (*42*). Concisely, adaptor proteins are characterised by the presence of four topological domains: an α-helical hairpin domain, a lipoyl domain, a β-barrel domain and a membrane proximal (MP)-domain. In general, PAPs exhibit well-defined modules; however, some structural domains are only shared within a subset of the family (*38*). The light shed into their structural organisation and

function suggests that these diverse, modular proteins are central players in the transport processes, mediating substrate recognition and selection, and intervening in the control of energy flow (*38*, *42*).

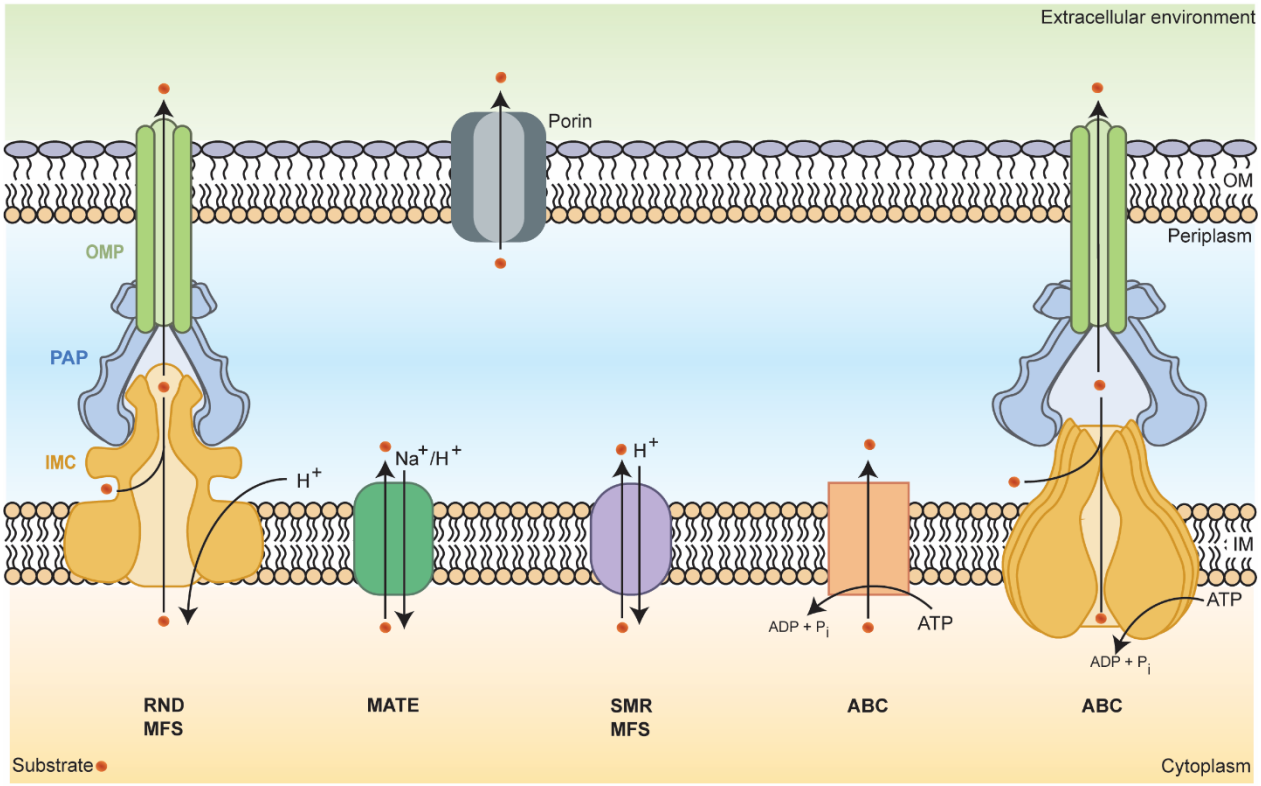


Figure 2: Schematic overview of multidrug (MDR) efflux pumps in Gram-negative bacteria.

MDR efflux pumps of Gram-negative bacteria are divided into five superfamilies according to the energy coupling mechanism, the transport mode and phylogenetic analysis. ABC superfamily proteins use ATP to transport diverse chemical compounds across the cellular inner membrane. MFS, SMR, and RND proteins function via an H^+ antiporte mechanism, whereas MATE proteins use both H^+ and Na^+ as energy source. RND, MFS and ABC transporters in particular can form multiprotein complexes that bridge the inner and outer membranes, accomplishing transport in a single step through an OMP belonging to the OEP family. Adapted from Delmar et *al.*, 2014 (*35*) and Anes et *al.*, 2015 (*36*).

1.3.1.2 TolC-dependent protein secretion

One of the most fundamental prokaryotic cell function is the transport of proteins from the cytoplasm to the cell surface, the extracellular environment, and/or other bacteria or eukaryotic cells in a process known as protein secretion. The exported proteins have key roles in the response of a bacterium to environmental cues and also in several physiological processes such as adhesion, pathogenicity, adaptation and survival (*34*). To surpass the two membranes and the periplasmic space enclosed within, bacteria have evolved large and complex multiprotein assemblies that either span the periplasm or establish two-step mechanisms employing periplasmic intermediate

steps (*34*, *43*). The TolC-dependent T1SS contrasts with these pathways because it bypasses the periplasm but requires only the OMP TolC, acting with an IM ABC-transporter and a PAP (*34*). For comparison purposes, a concise overview of the T1SS and two other protein secretion systems is presented in **Figure 3**.

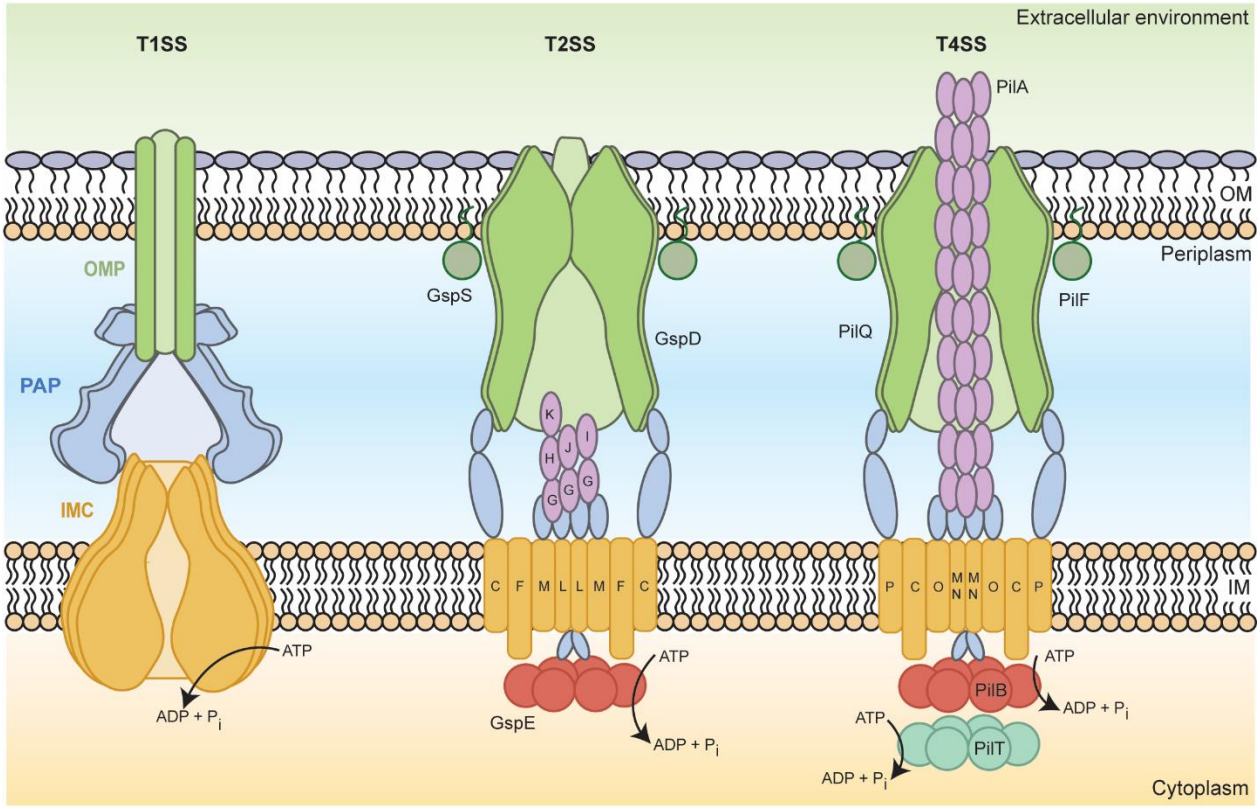


Figure 3: Schematic overview of the type I, type II and type IV protein secretion systems in Gram-negative bacteria.

The type I secretion system (T1SS) is a simple, tripartite system that facilitates the passage of proteins of various sizes across the cell envelope of Gram-negative bacteria. It consists of an ATP-binding cassette (ABC) transporter that constitutes the inner membrane (IM) component (IMC), a periplasmic adaptor protein (PAP) that bridges the IM and the outer membrane (OM), and an OM protein (OMP). Substrates are secreted in a single step. Contrastingly, the type II secretion system (T2SS) is a multicomponent machine (12-16 protein components) that uses a two-step mechanism for translocation. During the first step, the precursor effector protein is translocated across the IM by the Sec translocon or the twin-arginine-translocation (Tat) pathway. As an example, the general secretion pathway (Gsp) system is represented. The T2SS shows an evolutionary relationship with the type IV secretion system (T4SS), also known as the type IV pilus assembly machinery. The T4SS is a versatile system found in Gram-negative and Gram-positive bacteria that secretes a wide range of substrates, from single proteins to protein-protein and protein-DNA complexes. As an example, the assembly of the pilin polypeptide A (PilA) is schematised. Adapted from Costa et *al*., 2015 (*34*) and Korotkov et *al.*, 2012 (*44*).

The T1SS, sometimes referred to as ABC-dependent, has been found in a large number of Gram-negative bacteria, including pathogens of plants and animals (*45*). Substrates of this protein secretion system have an impressive range in size, from small 20 kDa proteins up to 900 kDa, and take part in different biological processes, such as nutrient acquisition (proteases and

lipases) and virulence (e.g. haemolysin) (*34*, *45*). In the T1SS, the multiprotein complex is typically assembled upon interaction of the ABC-transporter with a non-cleavable, C-terminal secretion signal that can be found in the protein, even though some substrates do not possess the C-terminal signal peptide (e.g. colicins and microcins) (*45*). ABC-transporters have been subdivided into three groups, based on the nature of their N-terminal sequences. The first class of ABC-transporters contains a C39 peptidase domain, which belongs to the papain superfamily structural motif, and is critical for recognising and cleaving the N-termini of substrates, as it happens with the colicin V of *E. coli*. On the other hand, ABC-transporters belonging to the second group possess a C39-like domain that lacks proteolytic activity and thus do not cleave its substrates. Many of the substrates recognised by this class of IMCs contain the characteristic repeats-in-toxin (RTX) motifs, rich in glycine and aspartic acid residues, which specifically bind calcium ions. Finally, the third group is characterised by the absence of any additional sequences in the N-terminal domain. In this case, substrates may or may not contain RTX motifs, but are smaller than those recognised by the second group of IMCs, and do present a secretion signal at their C-termini (*34*, *45*). Even though the adaptor proteins play a key role in these tripartite multiprotein complexes, they were the last component to be structurally characterised, greatly hindering knowledge about these proteins. Recently, the crystal structure of a part of the C-terminal periplasmic domain of HlyD, the archetypal PAP of the T1SS from *E. coli*, was solved and a RND-like interaction model with TolC was proposed, based on the structural similarities between the α-helical tip region of HlyD and PAPs in the RND efflux pumps (*46*).

1.3.1.3 TolC-dependent transport in cyanobacteria: a domain yet to be explored

TolC-dependent transport mechanisms have been largely explored in proteobacteria. In contrast, knowledge about these multiprotein assemblies in cyanobacteria remains scarce. The first cyanobacterial TolC homologue was identified in *Anabaena* sp. PCC 7120 (*47*). This homotrimeric OMP (HgdD, Alr2887) plays a role in the deposition of the glycolipid layer on the heterocyst envelope by interacting with the ATP-driven glycolipid efflux pump DevBCA (*47*–*49*). The function of HgdD as a major component of the metabolite export system was further demonstrated by proving its involvement in the translocation of antibiotics, exogenous compounds (e.g. ethidium bromide), proteins and siderophores (*47*, *50*–*53*). Subsequently, an effort has been made to identify and characterise the remaining components of this tripartite TolC-dependent transport system, and several proteins were identified (*50*, *52*, *53*). Similarly, an RND antiporter involved in the export of free fatty acids in *Synechococcus elongatus* PCC 7942 was recently identified, even though the TolC homologue has not yet been characterised (*54*).

Among all cyanobacterial strains, *Synechocystis* sp. PCC 6803 (hereafter referred to as *Synechocystis*) is one of the most commonly studied. *Synechocystis* is a unicellular, non-nitrogen-fixing cyanobacterium that exhibits versatile carbon metabolisms, growing under photoautotrophic, mixotrophic and heterotrophic conditions (*55*). This naturally competent organism was the first phototroph to have its genome fully sequenced (*56*, *57*) and annotated, being publicly available at CyanoBase (*58*). Despite its increasing relevance as a cell factory, not much is known about TolC-dependent export mechanisms in this cyanobacterium. In 2014, a TolC homologue (Slr1270) was identified in *Synechocystis* and a first glimpse of its secondary structure and function as an OMP provided (*59*). Afterwards, it was demonstrated that this TolC-like protein is involved in the export of proteins and other cellular metabolites, from structural elements of the cell wall (S-layer) to antibiotic resistance mechanisms (*60*). Its physiological relevance was also investigated and a mutant lacking a functional TolC protein displays an interesting set of phenotypic traits: increased activity of enzymes detoxifying reactive oxygen species (ROS), slower growth rates at low temperatures, impairment in biofilm formation, and increased release of outer membrane vesicles (OMV) when compared to the wild-type (*60*). Nonetheless, several questions remain unanswered: (i) what are the remaining components of these tripartite systems? (ii) how do these components interact and affect cell physiology? and (iii) which compounds are extruded by which transporter? A deeper understanding of these matters will certainly provide new valuable insights into TolC-dependent export mechanisms and lead the way to promising biotechnological applications.

1.4 Objectives

The main objective of this work is to disclose some of the fundamental mechanisms behind TolC-dependent transport systems in the unicellular cyanobacterium *Synechocystis* by addressing the following issues: (i) identify the genes that encode the putative IMCs and PAPs and characterise these proteins resorting to *in silico* tools; (ii) proceed to construct knockout mutants of these candidate genes with the aim of experimentally verify the function of the proteins they encode; (iii) ascertain which types of molecules, namely proteins, antibiotics, detergents and dyes, are preferentially exported by these putative transporter components.

2. MATERIALS AND METHODS

2 Materials and Methods

2.1 Organisms and standard growth conditions

The unicellular cyanobacterium *Synechocystis* sp. PCC 6803 wild-type was grown, unless stated otherwise, in BG11 medium (*61*), in Erlenmeyer flasks with orbital shaking (approximately 150 rpm) under a regimen of 16 h light (25 µmol photons m^{-2} s^{-1})/8 h dark at 25 °C. Knockout mutants were either maintained in solid BG11, containing 1 % (w/v) noble agar (Difco), 0.3% (w/v) sodium thiosulfate and kanamycin (10, 25, 50, or 100 µg mL^{-1}) or in liquid BG11 medium supplemented with 100 µg mL^{-1} of kanamycin.

For cloning purposes, *Escherichia coli* strain XL1-Blue (Stratagene/Agilent) was used. *E. coli* cells were propagated in solid LB medium (*62*) supplemented with 1.5% (w/v) agar and 100 µg mL^{-1} ampicillin and/or 50 µg mL^{-1} kanamycin, and grown overnight at 37 °C. *E. coli* transformants were cultured in selective liquid LB medium overnight at 37 °C on an orbital shaker (approximately 180 rpm).

2.2 Organisms growth measurements

Growth of cyanobacterial cultures and *E. coli* was regularly monitored by measuring the optical density at 730 nm (OD_{730}) or 600 nm (OD_{600}), respectively. Moreover, the chlorophyll *a* content was also adopted as a growth measurement (*63*). For that purpose, 1 mL of cyanobacterial culture suspension was harvested, centrifuged and the cell pellet suspended in 90% methanol. Samples were incubated overnight in a dark environment at 4 °C, centrifuged and absorbance at 663 nm of the supernatant was measured. An extinction coefficient of 12.7 was used to calculate the chlorophyll *a* concentration (*63*). In both methodologies, a Shimadzu UVmini-1240 spectrophotometer was used. Additionally, the optical density of cyanobacterial cultures grown in 96-well microplates was measured using a Synergy 2 multi-mode microplate reader (BioTek) at 790 nm (OD_{790}).

2.3 *In silico* identification of candidate genes

For the purpose of identifying putative components of translocase complexes related to TolC-dependent transport mechanisms in *Synechocystis*, protein sequences belonging to *E. coli* and assigned to the RND and the T1SS superfamilies were retrieved from UniProt: AcrA (P0AE06), AcrB (P31224), AcrE (P24180), AcrF (P24181), HlyB (P08716), HlyD (P09986). Moreover, the amino acid sequences of DevB (Alr3710) and DevC (Alr3711) from *Anabaena* sp. PCC 7120 were

retrieved from CyanoBase. These sequences were subjected to a tBLASTn search against the complete genome of *Synechocystis* with the e-value threshold set to 10 (*64*). To narrow down the list of candidate genes and increase the accuracy of the selection process, BLAST analyses were further complemented with additional *in silico* approaches, namely searches for protein conserved domains (*65*) and α-helical transmembrane segments (*66*). Additionally, the LipoP 1.0 Server was used to predict lipoproteins and discriminate between lipoprotein signal peptides, other signal peptides and N-terminal transmembrane helices (*67*).

2.4 Strategy for the construction of knockout mutants

2.4.1 Primer design

Selection of short nucleotide sequences (18 to 20 bp), flanking a region of approximately 500 bp, was performed resorting to Primer3 (*68*). For each target gene, two sets of specific oligonucleotides identified by the suffixes 5F/5R or 3F/3R were designed (**Table 1**). These primers aim to amplify two segments within the open reading frame (ORF) of interest in the genome of *Synechocystis*: the 5' and the 3' homologous regions. To introduce the necessary restriction sites for directional cloning, the restriction sequences recognised by XhoI and PstI were added to the 5F and 5R oligonucleotides, respectively; while the primers used to amplify the 3' homologous region possess the restriction sequences recognised by PstI (3F) and BamHI (3R), as illustrated in **Figure 4A**. In order to optimise cleavage close to the end of the fragment, three additional nucleotides were added immediately before the restriction sequence. The designed oligonucleotides were ordered to STAB Vida (Lisbon, Portugal).

2.4.2 PCR amplification, DNA purification and cloning

The specific set of oligonucleotides previously designed was used to amplify by PCR the 5' and 3' homologous regions of each target gene (2.6.2 Phusion DNA polymerase). PCR products were separated by agarose gel electrophoresis and DNA fragments corresponding to the desired amplicons were purified (2.10 DNA recovery and purification). Subsequently, the purified DNA fragments were digested with restriction endonucleases: the 5' homologous region was digested with XhoI and PstI, while the 3' homologous region was digested with PstI and BamHI (2.7 Digestion reactions). The digested DNA fragments were cloned into the pBluescript SK (+) vector (Agilent) digested with XhoI and BamHI (2.8 Ligation reactions), as illustrated in **Figure 4B**. After the ligation reaction, the assembled vector was introduced into competent *E. coli* XL1-Blue cells by transformation (2.13.1 *Escherichia coli* XL1-Blue transformation), plasmid DNA was retrieved (2.10 DNA recovery and purification) and the integrity of the homologous regions was verified by

Sanger sequencing (STAB VIDA, Lisbon). Afterwards, the kanamycin resistance cassette from the vector pUC4K (GE Healthcare) was amplified by PCR, resolved by agarose gel electrophoresis and purified. The purified kanamycin cassette was then digested with PstI and cloned into the previously obtained vectors (also digested with PstI), as depicted in **Figure 4C**. Finally, *E. coli* XL1-Blue cells were chemically transformed, plasmid DNA was extracted and confirmed by restriction analysis. The selected plasmids were then used to transform *Synechocystis* (2.13.2 *Synechocystis* sp. PCC 6803 transformation). To verify if the knockout mutants were fully segregated, genomic DNA was isolated (2.5 Isolation of genomic DNA) and a PCR analysis was performed (2.6.1 Taq DNA polymerase). A new set of primers was designed to evaluate chromosome segregation in the Δ*slr0454* mutant (**Table 1**), so that the difference between the size of the wild-type and the mutant amplicon was maximised.

Table 1: Oligonucleotides used in this work.

The unique restriction recognition sites added to the primers for directional plasmid cloning in *Escherichia coli* XL1-Blue are denoted in bold (XhoI), bold and italic (PstI) and bold and underlined (BamHI).

Target locus	Sequence (5'→3')	T_a (°C) [1]	Size (bp)	Reference
sll0141				
sll0141_5F	AAA**CTCGAG**GTCAAAGCCCAAACGGAGA	60	516	This work
sll0141_5R	AAA***CTGCAG***GCTCTGGTTTCTGCCTGTT			
sll0141_3F	AAA***CTGCAG***AGATCGAACAGCAGAGGGT	60	577	This work
sll0141_3R	AAA**GGATCC**CCTGCCACAATTTCCACCT			
sll0142				
sll0142_5F	AAA**CTCGAG**AACCAGTTGCCGGATGATC	60	596	This work
sll0142_5R	AAA***CTGCAG***TCTGTGCGTAACTCCTCCA			
sll0142_3F	AAA***CTGCAG***GCTTAATGGCGATCGTGGA	60	551	This work
sll0142_3R	AAA**GGATCC**TTTCACTCGCACATCCACC			
sll0180				
sll0180_5F	AAA**CTCGAG**GCTCCCAATTCCCTGTG	61	564	This work
sll0180_5R	AAA***CTGCAG***CACCCCCGTTTCCACTA			
sll0180_3F	AAA***CTGCAG***TAATTACCAAGCCGCCGTG	61	564	This work
sll0180_3R	AAA**GGATCC**TCGTGCCATCTTTGCCTT			
sll1053				
sll1053_5F	AAA**CTCGAG**GGTGGCTACTGGTGGTTTC	64	615	This work
sll1053_5R	AAA***CTGCAG***TTGTTGTAACCGTCGTTGG			
sll1053_3F	AAA***CTGCAG***AGTGGGCTACATTCCCGTA	64	512	This work
sll1053_3R	AAA**GGATCC**CCGTGATAACCCTGTCTCC			

Table 1: Continued

Target locus	Sequence (5'→3')	T_a (°C) [1]	Size (bp)	Reference
sll1180				
sll1180_5F	AAA**CTCGAG**CCACAGTTTGCCCAGTGAT	61	513	This work
sll1180_5R	AAA***CTGCAG***ACCGCAAATGGGCAAAGA			
sll1180_3F	AAA***CTGCAG***AAGAAACGCCCAAACCCA	61	586	This work
sll1180_3R	AAA**GGATCC**TCGGGAAAGCAGTTTGGTC			
sll1181				
sll1181_5F	AAA**CTCGAG**AGCCATTGTCATGCGTCA	61	567	This work
sll1181_5R	AAA***CTGCAG***AATTCAGCAGCGGTCACA			
sll1181_3F	AAA***CTGCAG***TGGCACGGTTACAGGTAGA	61	532	This work
sll1181_3R	AAA**GGATCC**TGCGCACGAAACCGATAT			
sll1481				
sll1481_5F	AAA**CTCGAG**ACCATAGCCTCCGTTGGAT	60	571	This work
sll1481_5R	AAA***CTGCAG***TGGCCGCTTGATTTACCTC			
sll1481_3F	AAA***CTGCAG***AACCATGGTGCTGGAAAGG	60	509	This work
sll1481_3R	AAA**GGATCC**CACTTTGGCGTTGCTCAAG			
sll1482				
sll1482_5F	AAA**CTCGAG**AAAATTCCCCTTGCCTGGC	60	570	This work
sll1482_5R	AAA***CTGCAG***ATTAAAGTCCCATCCGCCC			
sll1482_3F	AAA***CTGCAG***CAGTTCCATTGATGCGGGT	60	437	This work
sll1482_3R	AAA**GGATCC**TGGCCGTGGTCATAATCAG			
slr0369				
slr0369_5F	AAA**CTCGAG**GAAAACGATGACCAGGGCA	61	594	This work
slr0369_5R	AAA***CTGCAG***AGATGCCGTCACAAACAGG			
slr0369_3F	AAA***CTGCAG***AGACTGCTGGCTATTCCCA	61	601	This work
slr0369_3R	AAA**GGATCC**ATTGGCCTTTTCCCGATCC			
slr0454				
slr0454_5F	AAA**CTCGAG**TTTCCGGAAGTGAGCTTTC	64	506	This work
slr0454_5R	AAA***CTGCAG***AACACATCCTGGCCGTTAT			
slr0454_3F	AAA***CTGCAG***CTGGAGGAAAAGGTGATGG	64	603	This work
slr0454_3R	AAA**GGATCC**CAGAGCTAGGGGGAGCATA			
slr0454SegF	CATAACGGCCAGGATGTGTT	-	1320	This work
slr0454SegR	CAAAGCCATCACCTTTTCCT			
slr2131				
slr2131_5F	AAA**CTCGAG**ACATTTTGGAGCGGCAGA	61	526	This work
slr2131_5R	AAA***CTGCAG***TGGCGAGCAAGTCCAATT			
slr2131_3F	AAA***CTGCAG***ACGGGGTCAGCTAACTTTG	61	559	This work
slr2131_3R	AAA**GGATCC**ACCAACATCCCGCCAAAT			

Table 1: Continued

Target locus	Sequence (5'→3')	T_a (°C) [1]	Size (bp)	Reference
Kanamycin resistance cassette				
pUC4KF	AAA***CTGCA***GTGAGGTCTGCCTCGTGAAGAA	60	1228	(*60*)
pUC4KR	AAA***CTGCA***GAAAGCCACGTTGTGTCTCAAA			

[1] The annealing temperatures indicated refer to polymerase chain reactions carried out with the Phusion DNA polymerase (Thermo Fisher Scientific) (2.6.2 Phusion DNA polymerase). All amplifications performed with GoTaq DNA polymerase (Promega) were carried out at 59 °C (2.6.1 Taq DNA polymerase).

2.5 Isolation of genomic DNA

Cyanobacterial genomic DNA was extracted using the phenol-chloroform method, as previously described by Tamagnini *et al.,* 1997 (*69*). Concisely, *Synechocystis* cells were suspended in 10 mM Tris-HCl, pH 8.0, with 0.1 mM EDTA and disrupted by adding 0.3 g of acid-washed glass beads (425-600 µm diameter, Sigma-Aldrich), 25 µL of 10% sodium dodecyl sulphate (SDS), and 500 µL of phenol chloroform (1:1 [v/v] and vortexing at high speed). The liquid phases were separated by centrifugation at 14000 × *g* for 15 min, and the upper aqueous phase was extracted twice with an equal volume of chloroform. The DNA was precipitated with 1/10 volume of 3 M sodium acetate (pH 5.2) and 2.5 volumes of 100% ethanol at -20 °C for several hours before being washed, dried, and dissolved in water.

2.6 Polymerase chain reaction

2.6.1 Taq DNA polymerase

20 ng of genomic DNA or 4 ng of plasmid DNA were subjected to amplification in 20 µL reaction mixtures containing 0.25 µM of each primer (**Table 1**), 1× of Green GoTaq reaction buffer (Promega), 200 µM of each deoxynucleotide triphosphate (dNTP) (Thermo Fisher Scientific) and 0.1 units of GoTaq DNA polymerase (Promega). PCR amplification was performed in a thermal cycler (Bio-Rad), as described: an initial denaturation step at 95 °C for 2 min, followed by 35 cycles of 30 s at 95 °C, 30 s at 59 °C (**Table 1**), and 3 min at 72 °C, and a final extension step at 72 °C for 7 min. Negative controls (no template DNA) were included in all assays.

2.6.2 Phusion DNA polymerase

For the purpose of amplifying the 5' and 3' homologous regions and the kanamycin resistance cassette, each reaction mixture contained 20 ng of genomic DNA or 4 ng of plasmid DNA, respectively, as template, 0.5 µM of each primer, 1× of Phusion HF buffer (Thermo Fisher Scientific), 200 µM of each dNTP (Thermo Fisher Scientific) and 0.4 units of Phusion DNA

polymerase (Thermo Fisher Scientific) in a total volume reaction of 20 µL. All PCR reactions were carried out in a thermal cycler (Bio-Rad) according to the herein detailed profile: an initial denaturation step at 98 °C for 30 s, followed by 35 cycles of 10 s at 98 °C, 10 s at 60-64 °C (**Table 1**), and 15 s at 72 °C, and a final extension step at 72 °C for 5 min. Negative controls (no template DNA) were included in all assays.

2.7 Digestion reactions

All digestion reactions were performed according to the manufacturer's instructions (Thermo Fisher Scientific). For double digestions, the optimal enzymatic reaction buffer was selected using the DoubleDigest Calculator (Thermo Fisher Scientific). Reactions were carried out in a thermal cycler (Bio-Rad) for 3h at 37 °C and subsequently inactivated by incubation at 80 °C for 10 min.

2.8 Ligation reactions

The ligation reactions were carried out using the T4 DNA ligase (Thermo Fisher Scientific), as per the manufacturer's instructions. The insert:vector molar ratio used was 3:1 for approximately 100 ng of vector. Reaction mixtures (20 µL) were incubated overnight at 22 °C and a heat inactivation step at 65 °C for 10 min was performed before bacterial transformation.

2.9 Agarose gel electrophoresis

Electrophoretic analysis of nucleic acids was carried out on 0.8 to 1% (w/v) agarose gels (NZYTech) containing 1× Tris/Acetic acid/EDTA (TAE) buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH 8, Bio-Rad) and stained with either ethidium bromide (Bio-Rad) or GreenSafe Premium (NZYTech). GeneRuler 1 kb DNA Ladder (Thermo Fisher Scientific) was used as the molecular weight marker. Gels were visualised and photographed on a gel documentation system (Bio-Rad).

2.10 DNA recovery and purification

The NZYGelpure kit (NZYTech) was used to purify DNA from TAE agarose gels or for direct purification of PCR products or enzymatic reactions. Plasmid DNA recovery from recombinant *E. coli* strains was performed using the NZYMiniprep kit (NZYTech).

2.11 Purity and concentration of nucleic acids

The purity (A_{260}/A_{280} ratio) and concentration of nucleic acids was assessed using a NanoDrop ND-1000 UV-vis spectrophotometer (Thermo Fisher Scientific).

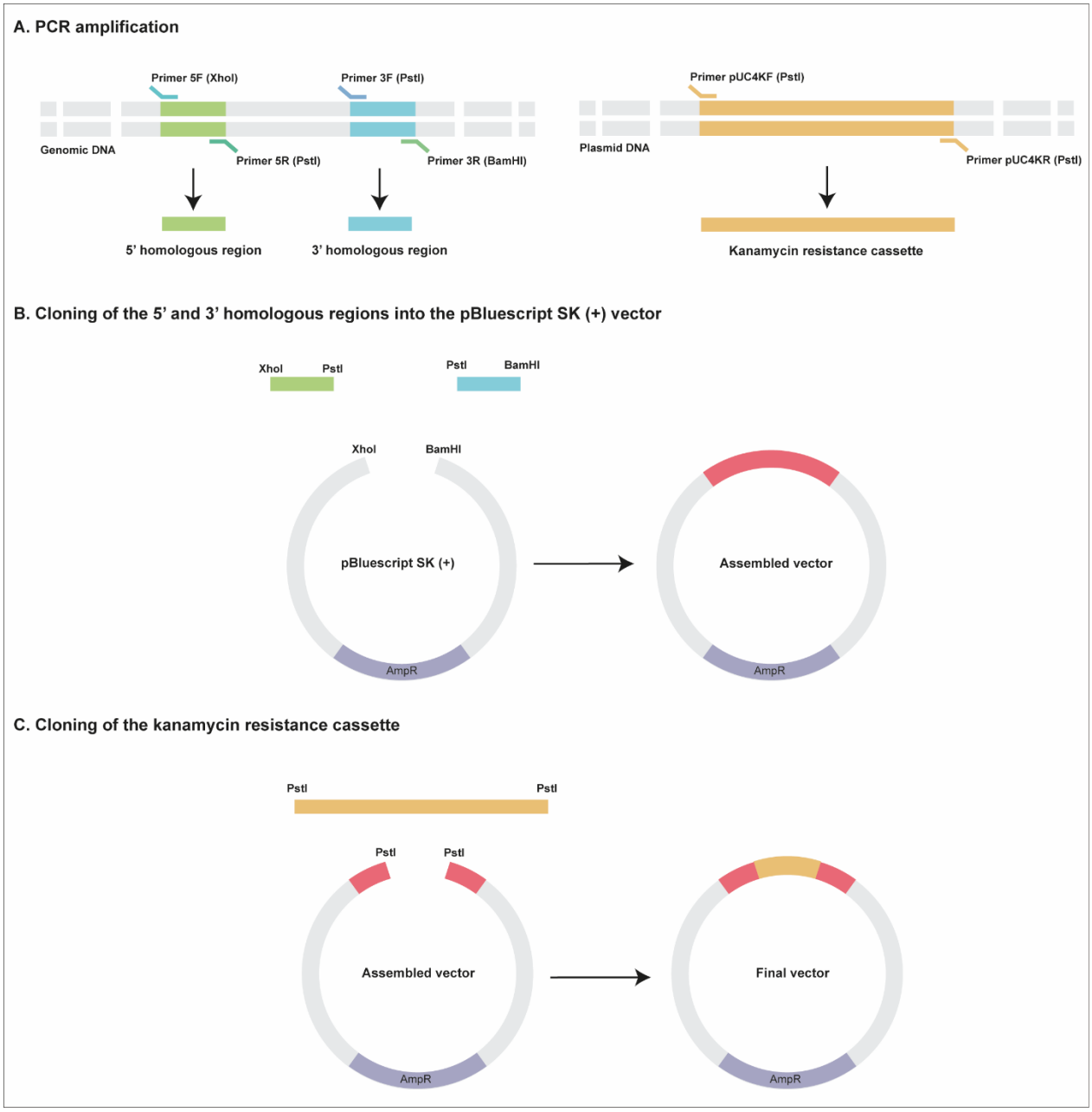


Figure 4: Strategy for the construction of knockout mutants in *Synechocystis* by insertional inactivation.

(A) For each target gene, two sets of specific oligonucleotides, primers 5F/5R and primers 3F/3R, were designed. The restriction sequences required for directional cloning and added to the oligonucleotides are indicated between brackets. This set of oligonucleotides was used to amplify two segments within the open reading frame (ORF) of interest in the genome of *Synechocystis*: the 5' and the 3' homologous regions. Additionally, the primers pUC4KF/pUC4KR were used to amplify the kanamycin resistance cassette from the pUC4K vector. (B) The 5' and 3' homologous regions were digested with XhoI and PstI and PstI and BamHI, respectively, and cloned into the pBluescript SK (+) vector digested with XhoI and BamHI. (C) The kanamycin resistance cassette was digested with PstI and cloned into the previously assembled vector digested with the same restriction endonuclease.

2.12 Chemical induction of competence

A variant of the Hanahan protocol (*70*) was employed to prepare competent cells of *E. coli* XL1-Blue. Briefly, super optimal broth (SOB) medium (0.5% (w/v) yeast extract, 2% (w/v) tryptone, 10 mM NaCl, 2.5 mM KCl, and 20 mM $MgSO_4$) was inoculated with one vial of seed stock and cells were grown overnight at 37 °C with shaking until they reached an OD_{600} of 0.5. Afterwards, cells were pelleted by centrifugation (10 min at 4000 × *g*, 4 °C), gently suspended in 80 mL of ice-cold CCMB80 buffer (10 mM KOAc pH 7.0, 80 mM $CaCl_2.2H_2O$, 20 mM $MnCl_2.4H_2O$, 10 mM $MgCl_2.6H_2O$, 10% glycerol, pH 6.4) and incubated on ice for 20 min. Cells were centrifuged again as described above, the supernatant was discarded and the pellet was suspended in 10 mL of ice-cold CCMB80 buffer. The OD_{600} of the cell suspension was then measured and adjusted to a final OD of 2.0. Finally, cells were aliquoted (100 µL) into microcentrifuge tubes and stored at -80 °C.

2.13 Bacterial transformation

2.13.1 *Escherichia coli* XL1-Blue transformation

Chemically competent *E. coli* XL1-Blue cells were transformed with plasmid DNA using the heat shock method (*71*), as follows: *E. coli* cells were gently thawed on ice for about 10 min, then the sample DNA (7 µL of a given ligation reaction or 1 µL of purified plasmid DNA) was added to an aliquot of competent cells and the mixture was incubated for 30 min; subsequently, cells were heat shocked at 42 °C for 90 s in a water bath, and once more incubated on ice for 2 min. Finally, warm LB medium was added to the mixture, and samples were incubated on an orbital shaker at approximately 200 rpm for 1 h at 37 °C to allow cells to recover. An adequate volume was then spread onto solid LB medium supplemented with 100 µg mL^{-1} ampicillin or with both 100 µg mL^{-1} ampicillin and 50 µg mL^{-1} kanamycin, followed by an overnight incubation at 37 °C. *E. coli* transformants were transferred to liquid LB medium supplemented with the suitable antibiotic and incubated overnight at 37 °C. Plasmid DNA was isolated and purified as described (2.10 DNA recovery and purification).

2.13.2 *Synechocystis* sp. PCC 6803 transformation

Natural transformation of *Synechocystis* was performed as described (*72*). *Synechocystis* was grown in liquid BG11 medium in glass gas washing bottles at 30 °C under a continuous light regime of 35 µmol photons m^{-2} s^{-1}. Once the culture reached its mid-exponential phase (OD_{730} of 0.6), cells were harvested by centrifugation at 4470 × *g* for 10 minutes at room temperature. The cell pellet was suspended in fresh BG11 medium so as to obtain a final OD_{730} of 2.60 nm and then mixed with exogenous DNA to a final concentration of 10 µg mL^{-1}. Subsequently, the mixture was

incubated at 30 °C in low light conditions for 6 h before plating on BG11 plates without selection marker, which where incubated in the same conditions for 24 h. Afterwards, the surface of the plates was washed with liquid BG11 medium, the cell suspension was centrifuged for 5 minutes at 1100 × *g*, and approximately 1/3 of the volume was spread onto BG11 plates supplemented with 10 µg mL^{-1} kanamycin. These plates were incubated at 25 °C in a 16 h light/8 h dark regimen (25 µmol photons m^{-2} s^{-1}) and colonies were observed after 7 to 15 days. To obtain a fully segregated mutant (all copies of the genome are mutated), transformants were exposed to increasing concentrations of kanamycin (25, 50, and 100 µg mL^{-1}). After approximately six streaking steps, a selection of kanamycin resistant colonies was picked and transferred into liquid BG11 medium supplemented with 100 µg mL^{-1}. Genomic DNA was isolated as described elsewhere (2.5 Isolation of genomic DNA) and mutants were assayed by PCR using gene-specific primers (2.6.1 Taq DNA polymerase).

2.14 Characterisation of the knockout mutants

In the present work, several *Synechocystis* strains (**Table 2**) were submitted to a set of experimental methodologies with the aim of studying the function of the proteins encoded by the insertionally inactivated genes through phenotypic analysis.

Table 2: Wild-type and mutant strains used in this work.

Strain designation	Genotype	Antibiotic selection	Reference
Wild-type	-	-	-
BSMN1 (Δ*tolC*)	*slr1270*::*aph*	Km	(*60*)
Δ*sll0141*	*sll0141*::*aph*	Km	This work
Δ*sll0142*	*sll0142*::*aph*	Km	This work
Δ*sll0180*	*sll0180*::*aph*	Km	This work
Δ*sll1053*	*sll1053*::*aph*	Km	This work
Δ*sll1180*	*sll1180*::*aph*	Km	This work
Δ*sll1181*	*sll1181*::*aph*	Km	This work
Δ*sll1481*	*sll1481*::*aph*	Km	This work
Δ*sll1482*	*sll1482*::*aph*	Km	This work
Δ*sll1481*Δ*sll1482*	*sll1481-sll1482*::*aph*	Km	This work
Δ*slr0369*	*slr0369*::*aph*	Km	This work
Δ*slr0454*	*slr0454*::*aph*	Km	This work
Δ*slr2131*	*slr2131*::*aph*	Km	This work

aph: aminoglycoside 3'-phosphotransferase kanamycin resistance protein; Km: kanamycin

Firstly, the role of the candidate proteins in the export of different exogenous compounds was investigated by analysing the tolerance of the wild-type strain and the knockout mutants towards different antimicrobial compounds (2.14.1 Microtiter plate assays). Additionally, an agar-based screening methodology was employed to evaluate the efflux capacity of *Synechocystis* strains to the fluorescent dye ethidium bromide, a known substrate of TolC-dependent transport mechanisms (2.14.2 Agar-based ethidium bromide screening assay). Moreover, the role of the candidate proteins in the secretion of proteins was studied by analysing the extracellular protein content present in the growth medium of the various strains (2.14.3 Exoproteome isolation and analysis). Finally, the phenotypic analysis was also complemented by investigating basic proteomic differences between the strains (2.14.4 Total protein extraction and analysis), and by negative staining electron microscopy observations of selected strains (2.14.5 Electron microscopy).

2.14.1 Microtiter plate assays

The tolerance of several *Synechocystis* strains (**Table 2**) to different antimicrobial compounds was evaluated. For that purpose, the wild-type and mutant strains were inoculated into liquid BG11 medium at an initial OD_{730} of 0.05. The cultures were incubated at 30 °C, with aeration, under a continuous light regime of 35 µmol photons m^{-2} s^{-1}. Cells were harvested as soon as they reached their mid-exponential growth phase (OD_{730} of approximately 0.6-0.8) and an appropriate volume was diluted to a final OD_{730} of 0.4. Aliquots (150 µL) were removed and pipetted into Greiner Bio-One 96-well, clear bottomed microplates. Subsequently, 50 µL of stock solutions of the antimicrobial compounds assayed were pipetted into the respective well (**Figure 5**) so as to obtain the concentrations indicated in **Table 3**.

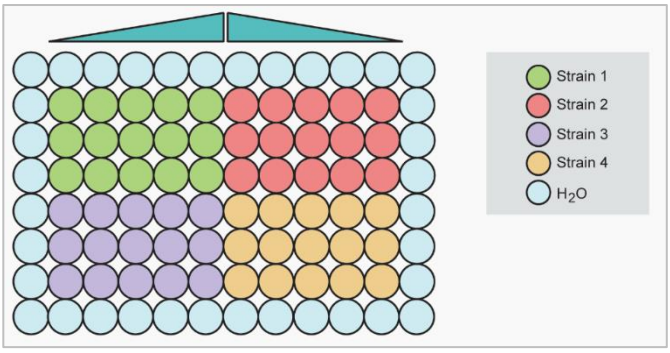


Figure 5: Schematic representation of the experimental set-up of the 96-well microplates.

In this experiment, the outermost wells were filled with 200 µL of water. For each bacterial strain, 3 technical replicates (indicated by the rows) were made. The highest concentration of the compound is located in the centre of the microplate, and the lowest concentration (absence of antimicrobial compound) is adjacently located to the wells filled with water. Microplates were incubated at 30 °C for 7 days under a 12h light (25 µmol photons m^{-2} s^{-1})/12h dark regimen.

In wells without antimicrobial compounds, 50 µL of distilled and deionized water were added instead. For each cyanobacterial strain, at least 3 biological and 9 technical replicates were performed. Microplates were incubated for 7 days at 30 °C, under a 12h light (25 µmol photons m^{-2} s^{-1}) /12h dark light regime on an orbital shaker (150 rpm). The OD_{790} was measured daily as previously described (2.2 Organisms growth measurements).

Table 3: List of antimicrobial compounds assayed in this work.

Chemical compound	Tested concentrations			
Detergents				
SDS (% w/v)	0.0025	0.0050	0.0125	0.0250
Triton X-100 (% v/v)	0.125	0.250	0.625	1.250
Antibiotics				
Chloramphenicol (µg mL^{-1})	0.050	0.100	0.250	0.500

2.14.2 Agar-based ethidium bromide screening assay

To analyse the efflux capacity of the wild-type and mutant strains (**Table 2**) to the fluorescent dye ethidium bromide, cells were inoculated into liquid BG11 (wild-type) or BG11 supplemented with 100 µg mL^{-1} kanamycin at an initial OD_{730} of 0.1. *Synechocystis* cultures were grown for 7 days at 30 °C in 25 mL Erlenmeyer flasks with orbital shaking (approximately 110 rpm) under a regimen of 12 h light (25 µmol photons m^{-2} s^{-1})/12 h dark. Subsequently, cells were harvested by centrifugation at 4300 × *g* for 5 minutes at room temperature. The cell pellet was suspended in liquid BG11 medium so as to obtain a final OD_{730} of 10, 5, 2.5, and 1. Afterwards, 10 µL of each dilution were pipetted onto solid BG11 supplemented with 0.5 µg ml^{-1} ethidium bromide, as depicted in **Figure 6**. Petri dishes were incubated at 25 °C for 24 h under a 16h light (25 µmol photons m^{-2} s^{-1})/8h dark regimen. Finally, petri dishes were visualised and photographed on a gel documentation system (Bio-Rad) and a densitometry analysis was carried out using ImageJ.

2.14.3 Data processing and statistical analysis

Data were statistically treated using the software package GraphPad Prism 6. Statistical analysis of the growth curves (OD_{790}) of the cyanobacterial strains grown in the absence of exogenous compounds was carried out using two-way ANOVA followed by Tukey's test. Growth inhibition (%) at 4 days was defined as (1 – $OD_{790\ \text{in the presence of exogenous compounds}}$ / $OD_{790\ \text{in the absence of exogenous compounds}}$) × 100 and the statistical significance analyses were achieved using one-way ANOVA followed by Tukey's test. Regarding the agar-based ethidium bromide screening assay, the relative intensity values were defined in function of the value registered for the wild-type and

the obtained data was compared by one-way Anova and the Tukey's multiple comparisons test. All significance values were expressed at the 95% confidence level ($p < 0.05$) or greater.

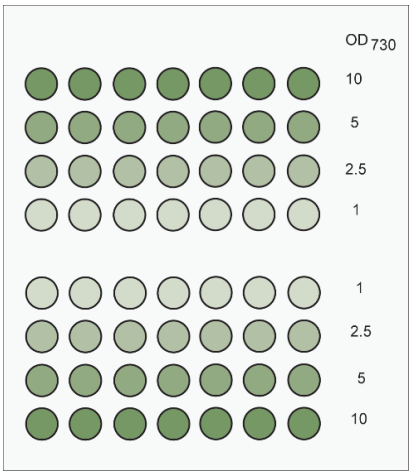


Figure 6: Schematic representation of the experimental set-up of the agar-based ethidium bromide screening assay.
In this assay, 10 µL of *Synechocystis* sp. PCC 6803 cells were pipetted onto solid BG11 medium supplemented with 0.5 µg mL^{-1} of ethidium bromide, as illustrated. Each column corresponds to a different bacterial strain and each row represents the optical density at 730 nm (OD_{730}) of the cyanobacterial culture.

2.14.4 Exoproteome isolation and analysis

The exoproteome was isolated according to the method described by Oliveira *et al.* (*73*). Briefly, *Synechocystis* was grown in liquid BG11 medium in gas washing bottles at 30 °C under a continuous light regime (35 µmol photons m^{-2} s^{-1}). Once the culture reached an OD_{730} of approximately 0.7-0.8 nm, 150 mL were collected by centrifugation at 4470 × *g* for 10 minutes at room temperature. The supernatant was filtered through 0.2 µm pore size filters (Fisherbrand) and further concentrated by centrifugation with Amicon Ultra-15 Centrifugal Filter units (Merck Milipore) with a nominal molecular weight limit of 3 kDa. Concentrated exoproteome samples were either used immediately or stored at -20 °C until further analysis. Samples were normalised for cell culture density (OD_{730}), volume of cell-free culture medium concentrated and concentration factor, before being separated by electrophoresis on gradient 4-15% (w/v) (Bio-Rad) SDS-polyacrylamide gels and visualised with colloidal Coomassie Brilliant Blue (Sigma) or with the Glycoprotein Staining Kit (Pierce). NZYColour Protein Marker II (NZYTech) was used as the molecular weight marker.

2.14.5 Total protein extraction and analysis

Synechocystis was grown in liquid BG11 medium at 30 °C in 25 mL Erlenmeyer flasks with orbital shaking (approximately 110 rpm) under a regimen of 12 h light (25 µmol photons m^{-2} s^{-1})/12 h dark. Once the culture reached a final OD_{730} of approximately 2.0, cells were harvested by centrifugation at 4470 × *g* for 10 minutes at 4 °C. Pellets were suspended in protein extraction

buffer [10 mM HEPES, 0.5 % (v/v) Triton X-100, 10 mm EDTA, 2 mm DTT, pH 8.0, supplemented with a protease inhibitor cocktail (Roche)]. Cells were sonicated (Branson Sonifier 250) for 5 cycles (30 s at power 3, 30 s rest, 50% duty cycle) and incubated on ice. Subsequently, samples were centrifuged at 11000 × *g* for 10 minutes at 4 °C, the supernatant was transferred into a new microcentrifuge tube and stored at -20 °C until further analysis. Total protein concentration was determined using the BCA method (Thermo Fisher Scientific), with bovine serum albumin as standard. Total protein samples were electrophoresed on 8% (w/v) SDS-polyacrylamide gels (29:1 acrylamide:bisacrylamide) and stained with colloidal Coomassie Brilliant Blue (Sigma). NZYColour Protein Marker II (NZYTech) was used as the molecular weight marker.

2.14.6 MALDI TOF/TOF mass spectrometry

Coomassie Blue stained bands containing peptides of interest were excised from the SDS-polyacrylamide gel, subjected to in-gel trypsin digestion and the peptide mass fingerprint was further analysed on a 4700 MALDI TOF/TOF (matrix-assisted laser desorption/ionization tandem time-of-flight mass spectrometry) proteomics analyser (Applied Biosystems).

2.14.7 Electron microscopy

Synechocystis cells were grown in liquid BG11 medium in glass gas washing bottles at 30 °C under a continuous light regime (35 µmol photons m^{-2} s^{-1}) up to an OD_{730} of 1.0. For negative staining transmission electron microscopy (TEM), 10 µL of cyanobacterial culture were mounted on formvar/carbon film coated with mesh nickel grids (Electron Microscopy Sciences) and left standing for 2 min. The liquid in excess was removed with paper filter, and 10 µL of 1% uranyl acetate were added onto the grids and left standing for 10 s, after which liquid in excess was removed with paper filter. Visualisation was carried out on a Jeol JEM-1400 transmission electron microscope.

3. RESULTS AND DISCUSSION

3 Results and Discussion

3.1 *In silico* identification of candidate genes and bioinformatic analysis of the proteins they encode

In spite of the current knowledge about TolC-dependent export mechanisms in Gram-negative bacteria, remarkably little is known about these systems in cyanobacteria, namely the molecular nature of the components involved. Previously, a TolC homologue (Slr1270) was identified in *Synechocystis* and its role as a major component of the metabolite export system was investigated, providing strong evidence that Slr1270 is a member of the OEP family (*59*, *60*). Proteins of the OEP family function together with a primary IM transporter and a PAP, establishing a multiprotein complex that allows the export of a broad range of substrates across the cell envelope in a single energy-coupled step (*32*). However, it should be noted that TolC itself is a rather useless protein, since its function relies on the transient interaction with the IM translocase complex which mediates substrate recognition and energises the export process (*31*). However, in *Synechocystis*, the IM translocase partners (IMC and PAP) have not been identified yet, resulting in a lack of knowledge about their biochemical relevance for the system's functionality. Therefore, one of the aims of this study was to employ computational methods to address this issue and identify members of the IM translocase complexes in this unicellular cyanobacterium.

In a first approach, protein sequences assigned to two of the most studied TolC-dependent export systems, the T1SS and the RND efflux pumps, were retrieved from UniProt. In *E. coli*, HlyBD is the only multiprotein complex proved to be involved in the T1SS (*34*), whereas six RND IM translocase complexes have been described: AcrAB, AcrAD, AcrEF, MdtAB, MdtEF, and CusCFBA (*36*). AcrAB is the archetypal RND efflux pump from *E. coli* and has been thoroughly studied and described (*74*–*77*). In contrast, knowledge about AcrEF is scarcer, despite its high homology to AcrAB (*78*, *79*). Additionally, the amino acid sequences of DevB and DevC, two components of a cyanobacterial TolC-dependent glycolipid efflux pump, were retrieved from CyanoBase. Subsequently, the Basic Local Alignment Search Tool (BLAST) available at CyanoBase was used to perform a tBLASTn search against the complete genome of *Synechocystis* (*64*). The BLAST output for each query sequence was curated and only the hits presenting the lowest *e*-values were selected, as shown in **Table 4**.

Table 4: Output of the tBLASTn searches performed against the amino acid sequences derived from the complete genome of *Synechocystis* sp. PCC 6803.

Query protein	ORF	Description	Score (bits)	*e*-value	Identity (%)
Periplasmic adaptor proteins					
AcrA (P0AE06)	*sll1053*	Hypothetical protein	62	2×10^{-11}	26
	sll0180	Hypothetical protein	58	5×10^{-10}	24
	sll0141	Hypothetical protein	42	3×10^{-5}	23
AcrE (P24180)	*sll1053*	Hypothetical protein	63	1×10^{-11}	27
	sll0180	Hypothetical protein	47	8×10^{-7}	23
	sll0141	Hypothetical protein	38	7×10^{-4}	23
HlyD (P09986)	*sll1181*	Similar to haemolysin secretion protein	75	3×10^{-15}	33
	sll1481	ABC-transporter membrane fusion protein	37	0.002	24
	sll0141	Hypothetical protein	34	0.021	21
DevB (Alr3710)	*sll1481*	ABC-transporter membrane fusion protein	279	3×10^{-90}	40
	slr1207	Hypothetical protein	55	5×10^{-9}	24
	sll0141	Hypothetical protein	54	7×10^{-9}	25
Inner membrane transporters					
AcrB (P31224)	*slr2131*	RND multidrug efflux transporter	786	0	40
	slr0369	RND multidrug efflux transporter	664	0	37
	sll0142	Probable cation efflux system protein	259	6×10^{-73}	24
	slr0794	Probable cation efflux system protein	233	4×10^{-64}	24
	slr6043	Probable cation efflux system protein	201	5×10^{-54}	23
	slr0454	RND multidrug efflux transporter	128	4×10^{-31}	31
AcrF (P24181)	*slr2131*	RND multidrug efflux transporter	792	0	42
	slr0369	RND multidrug efflux transporter	658	0	23
	sll0142	Probable cation efflux system protein	283	3×10^{-81}	37
HlyB (P08716)	*sll1180*	Toxin secretion ABC transporter ATP-binding protein	543	0	45
	sll1276	ATP-binding protein of ABC transporter	248	8×10^{-75}	32
	sll0484	ATP-binding protein of ABC transporter	100	2×10^{-24}	34
DevC (Alr3711)	*sll1482*	ABC transporter permease protein	430	1×10^{-151}	54
	slr0594	Hypothetical protein	37	0.002	25

UniProt or CyanoBase identifiers are indicated between brackets. Accessed on 21st September 2015.

Even though BLAST is frequently used to identify homologous genes/proteins, the alignments produced usually correspond to isolated regions of similarity, providing no indication about how they are structured in the gene/protein model (*64*). Consequently, some BLAST hits retrieved from such analysis result from the high homology level that highly-conserved, yet highly-frequent protein domains, generate and thus need to be further validated (e.g. HlyB harbours an ATP-binding cassette domain – ABC-type, which is highly frequent in the *Synechocystis* inferred

proteome). One of the most straightforward ways to do so is to gather information about the protein encoded by the candidate gene. Accordingly, the amino acid sequence encoded by each candidate gene was retrieved from CyanoBase and subjected to a conserved domain search through the NCBI's Conserved Domain Database (CDD), as these conserved regions of a protein often convey distinct function (*65*). Even though the majority of the protein repertoire in analysis is composed of multi-domain proteins, the approach undertaken in the present work categorised the candidate proteins according to the best domain match (**Figure 7**). Thus, proteins were classified either as PAPs or IM transporters and then subdivided into three distinct categories: the RND, the T1SS or the Dev. Subsequently, this classification was further refined by analysing the protein domains assignments at the superfamily level, depicted in **Figure 7**, and the absence/presence (and in the latter case, the number) of predicted transmembrane helices (**Figure 8**, **Figure 9**, and **Figure 10**).

The putative adaptor proteins identified herein are characterised, despite of their subcategory, by the presence of up to two conserved domains that belong to the biotinyl/lipoyl superfamily. It is known that bacterial adaptor proteins possess four topological domains, among them a lipoyl domain which is located adjacently to the hairpin and its helical extension (1.3.1.1 TolC-dependent multidrug efflux) (*38*). The exact functional role of this domain has not yet been established, but mutational analyses suggests it has a role in stabilizing the complex assembly either by interaction with the IM transporter or by self-association (*38*).

Regarding the putative IM transporters of the RND superfamily, it is possible to observe the existence of one (Slr0369, Slr0454, Slr2131) or two (Sll0142) mycobacterial membrane protein large domains (MMPL). In *Mycobacterium* species, MmpL transporters are dedicated to the export of mycobacterial lipids for cell wall synthesis and belong to a subfamily within the diverse RND superfamily (*80*). However, the precise function of this domain in Gram-negative RND transporters has not yet been disclosed. On the other hand, ABC transporters have a characteristic architecture that consists of two domains: a transmembrane domain, embedded in the membrane bilayer, and an ABC (or nucleotide-binding domain, NBD), located in the cytoplasm (*81*). Likewise, an ABC transmembrane (ABC membrane), and a NBD (ABC transporter) regions were identified in the putative T1SS IM transporter Sll1180. Additionally, a peptidase C39-like domain was also recognised. This domain is present in HlyB, and it has been postulated that intervenes in the initial recognition events of the secretion mechanism (*82*). Finally, Sll1482 was categorised as homologous to the IM transporter DevC. In the C-terminal region of this protein is located an FtsX-like permease domain. Interestingly, this domain is also present on the LolC transmembrane

protein of *E. coli*, as part of an ATP-dependent transport complex LolCDE responsible for the release of lipoproteins targeted to the OM from the IM (*83*).

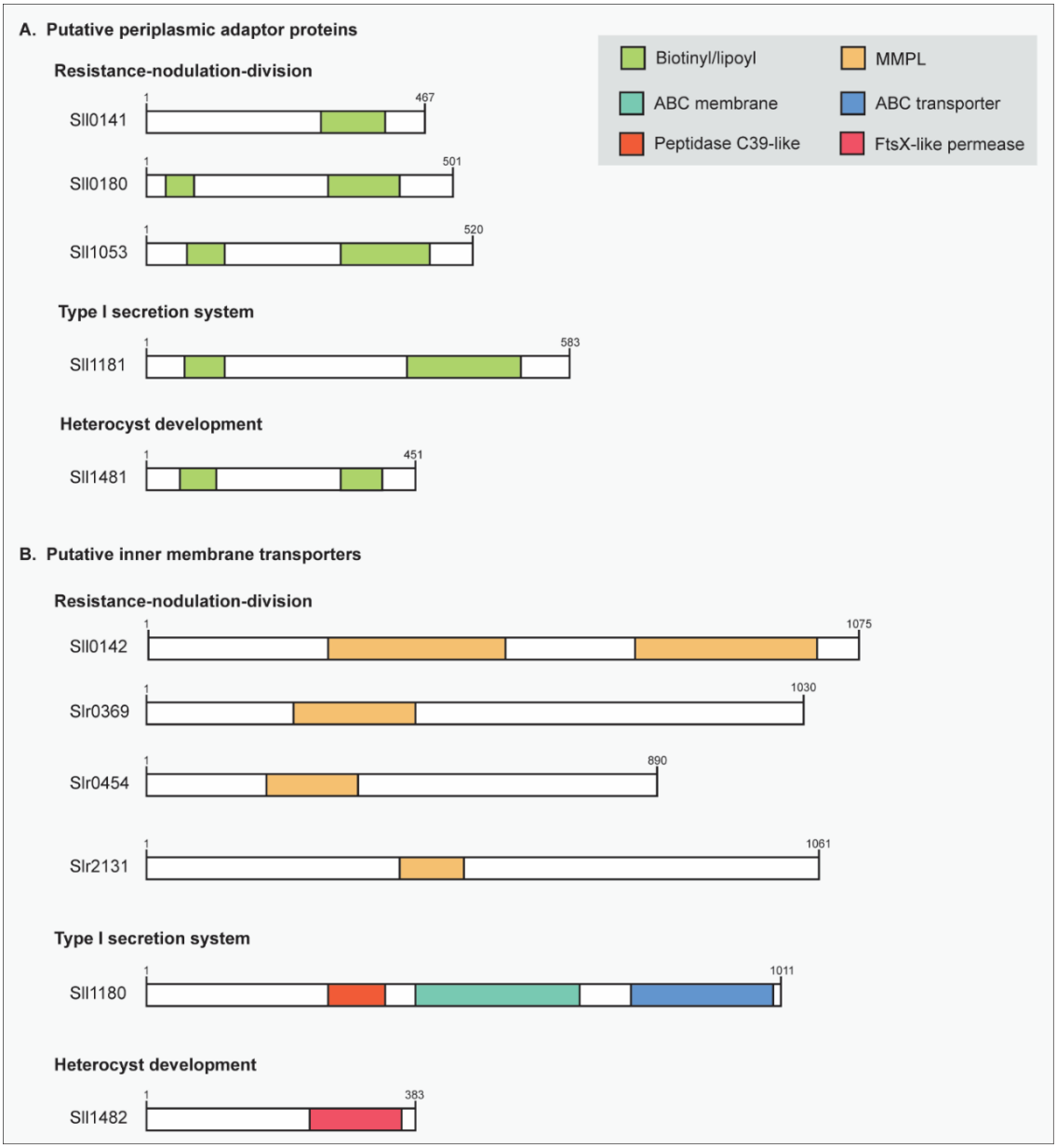


Figure 7: Conserved protein domains, closing in on the biological function of the putative candidate proteins.

The putative inner membrane translocase partners of TolC are classified as periplasmic adaptor proteins (A) or inner membrane transporters (B), according to the best domain hit. Within each major group, proteins fall into three distinct subcategories: the resistance-nodulation-division (RND), the type I secretion system (T1SS) and the heterocyst development (Dev). Proteins are depicted as white rectangles scaled to the length of the primary sequence. The organisation of the structural domains at the superfamily level is specified by different colours: MMPL, mycobacterial membrane protein large; ABC membrane, ATP-binding cassette membrane; ABC transporter, catalytic core of the nucleotide-binding domain. This functional annotation is based on searches through the NCBI's Conserved Domain Database (CDD) (*65*).

In TolC-dependent export mechanisms, the IM transporter is an integral membrane protein containing membrane-spanning α-helical domains, whereas the adaptor protein might be bound to the IM by a lipid anchor or possess an α-helical segment in its N-terminal region (*38*). Thus, the transmembrane topology of the putative candidate proteins was predicted resorting to the online tool PHOBIUS (*66*) and the obtained results are presented in **Figure 8**, **Figure 9**, and **Figure 10**.

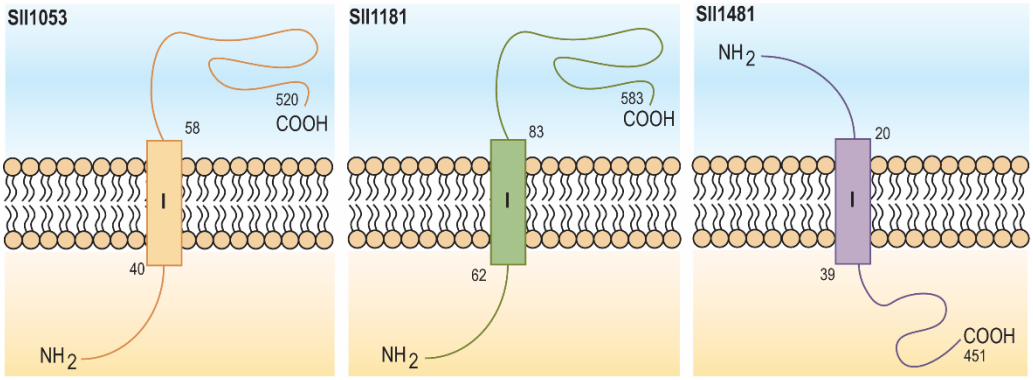


Figure 8: Schematic view of the predicted transmembrane topology of Sll1053, Sll1181 and Sll1481.

The putative periplasmic adaptor proteins Sll1053, Sll1181 and Sll1481 are predicted to form single-pass membrane proteins. The transmembrane helix is located in the N-terminal region of the protein. All predictions were carried out with PHOBIUS (*66*).

Usually, adaptor proteins complexed with transporters belonging to the ABC or MFS superfamily have a transmembrane helix, which connects the periplasmic domain with a small cytoplasmic domain (*31*, *38*). Out of all the putative PAPs identified, only Sll1181 and Sll1481 are expected to interact with ABC transporters. Accordingly, these putative accessory proteins are predicted to form single-pass membrane proteins (**Figure 8**). In both cases, the hypothetical α-helical transmembrane segment is located in the N-terminal region of the protein; however, in Sll1181 the N-terminal region seems to be cytoplasmic, whereas in Sll1481 this region appears to be located in the periplasmic space. Alternatively, PAPs interacting with RND transporters are typically bound to the IM by an N-terminal lipid anchor (*31*, *84*). Nonetheless, some exceptions are known, for instance, the accessory protein EmrA possesses an N-terminal transmembrane helix, even though it shares similarities with PAPs related to RND efflux pumps (*31*). In the present work, searches for protein conserved domains classified Sll0141, Sll0180 and Sll1053 as the PAP subunit of RND IM translocase complexes. Hypothetical transmembrane-spanning segments were not detected in Sll0141 and Sll0180. However, Sll1053 is predicted to form a single-pass integral membrane protein, which has its N-terminal region located in the cytoplasm. Alternatively, the amino acid sequences of Sll0141 and Sll0180 were submitted to the LipoP 1.0 Server, which predicts lipoprotein signal peptides in Gram-negative bacteria (*67*). An N-terminal lipoprotein signal

sequence was recognised in both proteins with a cleavage site by the signal peptidase II (SPaseII) located between the amino acids 23 and 24 in Sll0180 and 26 and 27 in Sll0141.

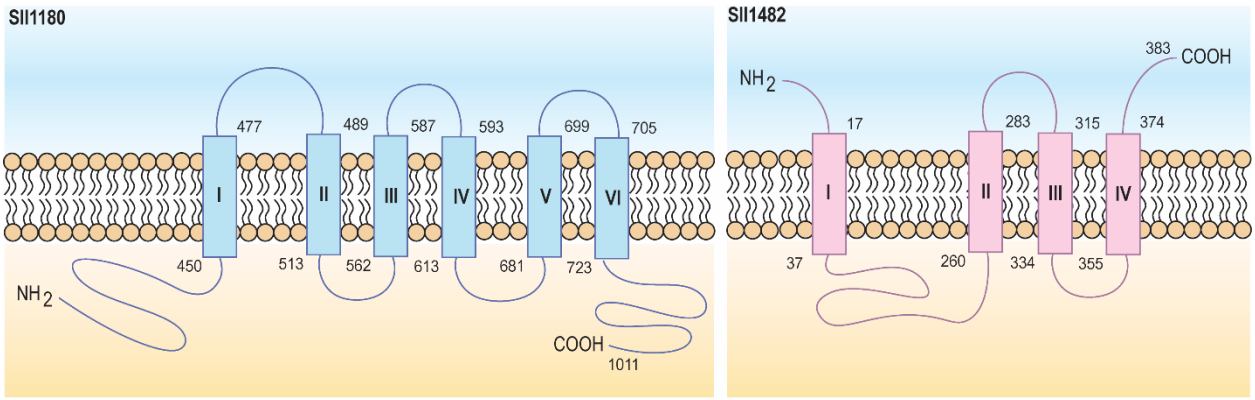


Figure 9: Schematic view of the predicted transmembrane topology of Sll1180 and Sll1482.

The putative inner membrane transporters Sll1180 and Sll1482 are predicted to form multi-pass membrane proteins with six and four α-helical transmembrane segments, respectively. All predictions were carried out with PHOBIUS (*66*).

Sll1180, the putative ABC transporter presumably involved in the T1SS mechanism, is predicted to form a multi-pass membrane protein with six transmembrane α-helical segments (**Figure 9**). The membrane domain of HlyB was experimentally determined to have eight membrane-spanning regions (*85*), even though *in silico* tools usually predict a lower number of transmembrane segments (four to seven). Unlike RND efflux pumps, the 3D structure of the T1SS ABC transporters remains elusive, as only the structure of nucleotide binding domain has been solved (*86*). In this case, the challenging nature of membrane proteins has been hindering the much necessary progress to thoroughly understand their topological organisation. Nevertheless, it is known that the ATP-binding cassette of HlyB is located in the cytoplasmic C-terminal domain of the protein, as depicted in **Figure 7** for Sll1180. Furthermore, it has been hypothesized that the N-terminal cytoplasmic loop is involved in substrate recognition and specificity (*87*).

The structure of DevC has not yet been solved. However, its transmembrane topology was also predicted with the purpose of establishing a point of comparison between this protein and its *Synechocystis* homologue (Sll1482). Thus, DevC is predicted to form a multi-pass protein wherein four transmembrane α-helical segments were recognised. A short N-terminal region was predicted as well as a periplasmic loop of approximately 200 residues located between the transmembrane segments I and II (data not shown). Likewise, the DevC homologue Sll1482 was computationally determined to have four α-helices (**Figure 9**). Nevertheless, bioinformatics analyses suggest that

in this cyanobacterial protein the N-terminal region may be periplasmic, while the cytoplasmic loop is possibly placed between the transmembrane segments I and II.

IM transporters of the RND superfamily are usually characterised by the presence of twelve transmembrane-spanning segments and two large, hydrophilic periplasmic loops, which constitute one of the most distinctive features of this class of transporters (*88*, *89*). In comparison, other IM transporters, such as EmrB (*90*) or HlyB (*91*), which are also components of tripartite efflux systems, do not contain large periplasmic loops. It has been reported that the two periplasmic loops are involved in substrate recognition and specificity (*92*, *93*). A mutational analysis carried out in *mexD*, which encodes the IM transporter of the MexCD–OprJ efflux pump from *Pseudomonas aeruginosa*, revealed that all the mutations capable of altering the substrate specificity of this pump were mapped to the periplasmic loops of MexD (*92*). Furthermore, a chimeric study between AcrB and AcrD showed that the replacement of the two large periplasmic loops of AcrD with the equivalent loops of AcrB altered the substrate range of AcrD to a broader one, representative of AcrB. On the contrary, AcrB chimeras containing both loops of AcrD conferred resistance only to the typical substrates of AcrD (*93*). *In silico* predictions identified twelve transmembrane segments in the putative RND efflux pumps Sll0142, Slr0454 and Slr2131, as illustrated in **Figure 10**. Moreover, it is possible to observe that these putative RND transporters exhibit the two distinctive periplasmic loops, composed by approximately 300 residues, and located between the transmembrane segments I and II and the transmembrane segments VII and VIII. Slr0369, on the other hand, is predicted to form a multipass membrane protein with eleven α-helical transmembrane segments. In this protein, the N-terminal region seems to be periplasmic and larger (approximately 300 residues) than in other RND transporters, which might suggest a slightly unusual topological organisation.

The computational methods used herein provided important glimpses into the possible biological role of the proteins encoded by the candidate genes as well as into their membrane topology, supporting their presumptive involvement in TolC-dependent export mechanisms. However, *in silico* predictions are based on experimental data obtained from other organisms. This means that inferences regarding the T1SS or the Dev are weaker than those related to RND efflux pumps, as the latter system has been thoroughly studied in organisms such as *E. coli* and *P. aeruginosa*, and the structures of many components have been solved. Besides its role in the identification and classification refinement of candidate genes/proteins, bioinformatics tools also provide an important theoretical basis on which experimental methodologies can be grounded with the aim of addressing their function, and mechanisms of substrate recognition and specificity.

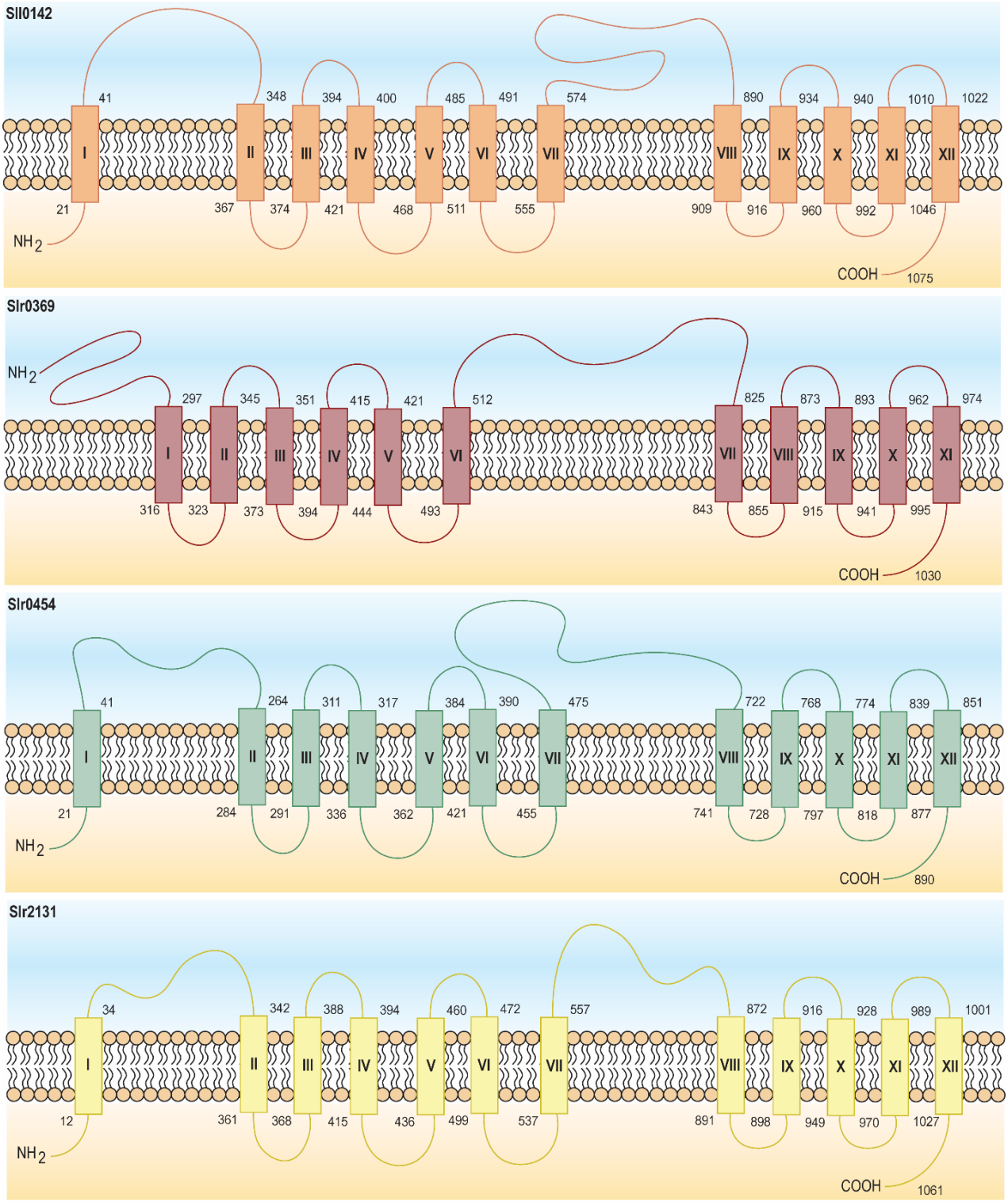


Figure 10: Schematic view of the predicted transmembrane topology of Sll0142, Slr0369, Slr0454 and Slr2131.

The putative inner membrane transporters of the RND superfamily Sll0142, Slr0369, Slr0454 and Slr2131 are multipass α-helical membrane proteins characterised by the presence of up to twelve transmembrane segments. All predictions were carried out with PHOBIUS (*66*).

3.2 Construction of knockout mutants

In the present work, bioinformatics tools revealed to be highly valuable to identify putative candidate proteins involved in TolC-dependent efflux systems in *Synechocystis*. However, the candidate proteins identified have particular aspects that diverge from the *E. coli* model proteins; therefore, it became imperative to validate the bioinformatics screening. For that purpose, knockout mutants of *Synechocystis* were constructed by insertional inactivation, as detailed in the Materials and Methods section (2.4 Strategy for the construction of knockout mutants). *Synechocystis* wild-type cells were transformed with the obtained DNA constructs by double homologous recombination. Mutant selection was carried out in solid BG11 medium supplemented with increasing concentrations of the selection marker, the antibiotic kanamycin. Since *Synechocystis* contains about 12 copies of the chromosome per cell (*94*), a PCR analysis was carried out to verify that homozygosity was obtained (**Table 5** and **Figure 11**).

Table 5: Amplicon size (bp) of the PCR products resulting from the amplification of genomic DNA of *Synechocystis* wild-type and mutant strains with gene-specific oligonucleotides.

Target locus	Wild-type	Mutant	Target locus	Wild-type	Mutant
sll0141	1211	2320	***sll1481***	1270	2307
sll0142	1998	2374	***sll1482***	2321	2234
sll0180	1285	2355	***sll1481 sll1482***	2362	2235
sll1053	1330	2354	***slr0369***	1789	2422
sll1180	1923	2326	***slr0454***	1320	1254
sll1181	1295	2326	***slr2131***	2385	2336

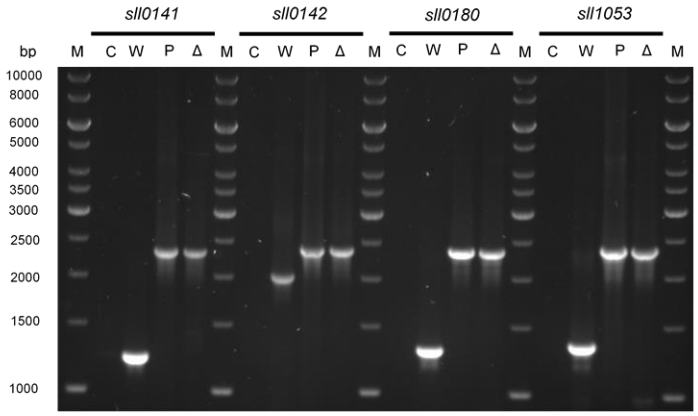


Figure 11: Assessment of the segregation of *Synechocystis* knockout mutants by PCR amplification of the target region.

PCR reactions were carried out using gene-specific oligonucleotides identified by the suffixes 5F/3R (**Table 1**), which flank the 5' and 3' homologous regions. C, negative control. W, amplification of genomic DNA isolated from *Synechocystis* wild-type. P, amplification of the purified plasmid DNA used to transform *Synechocystis* (positive control). Δ, amplification of genomic DNA extracted from *Synechocystis* knockout mutants. M, GeneRuler 1 kb DNA ladder (Thermo Fisher Scientific). The expected amplicon size (bp) is indicated in **Table 5**.

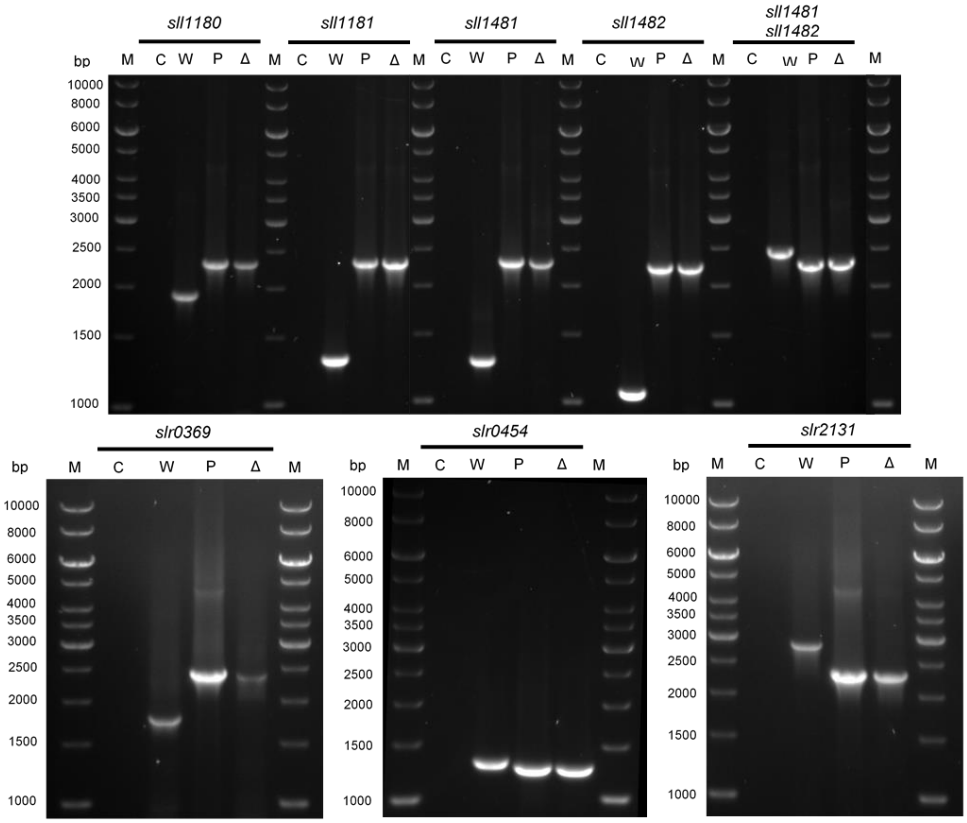


Figure 11: Continued.

PCR reactions were carried out using gene-specific oligonucleotides identified by the suffixes 5F/3R (**Table 1**), which flank the 5' and 3' homologous regions. Alternatively, a new set of gene-specific oligonucleotides was designed to evaluate chromosome segregation in the Δ*slr0454* mutant (**Table 1**), so that the difference between the size of the wild-type and the mutant amplicon was maximised. C, negative control. W, amplification of genomic DNA isolated from *Synechocystis* wild-type. P, amplification of the purified plasmid DNA used to transform *Synechocystis* (positive control). Δ, amplification of genomic DNA extracted from *Synechocystis* knockout mutants. M, GeneRuler 1 kb DNA ladder (Thermo Fisher Scientific). The expected amplicon size (bp) is indicated in **Table 5**.

From the figure above it is possible to observe that the knockout mutants constructed are fully segregated, indicating that the proteins encoded are not essential for cell survival under the standard laboratory conditions used. Subsequently, the knockout mutants were subjected to different experimental assays with the aim of characterising protein function by phenotypic analysis.

3.3 Growth assessment of the *Synechocystis* knockout mutants

In this work, eleven genes that presumably encode subunits of the IM translocase partners of TolC-dependent export mechanisms were inactivated in the genome of *Synechocystis*. To determine whether the proteins encoded by the disrupted genes have any effect on cell growth or viability, the wild-type and mutant strains were cultivated under the conditions stated in **Figure 12**, and cell growth (OD_{790}) was monitored daily for 4 days

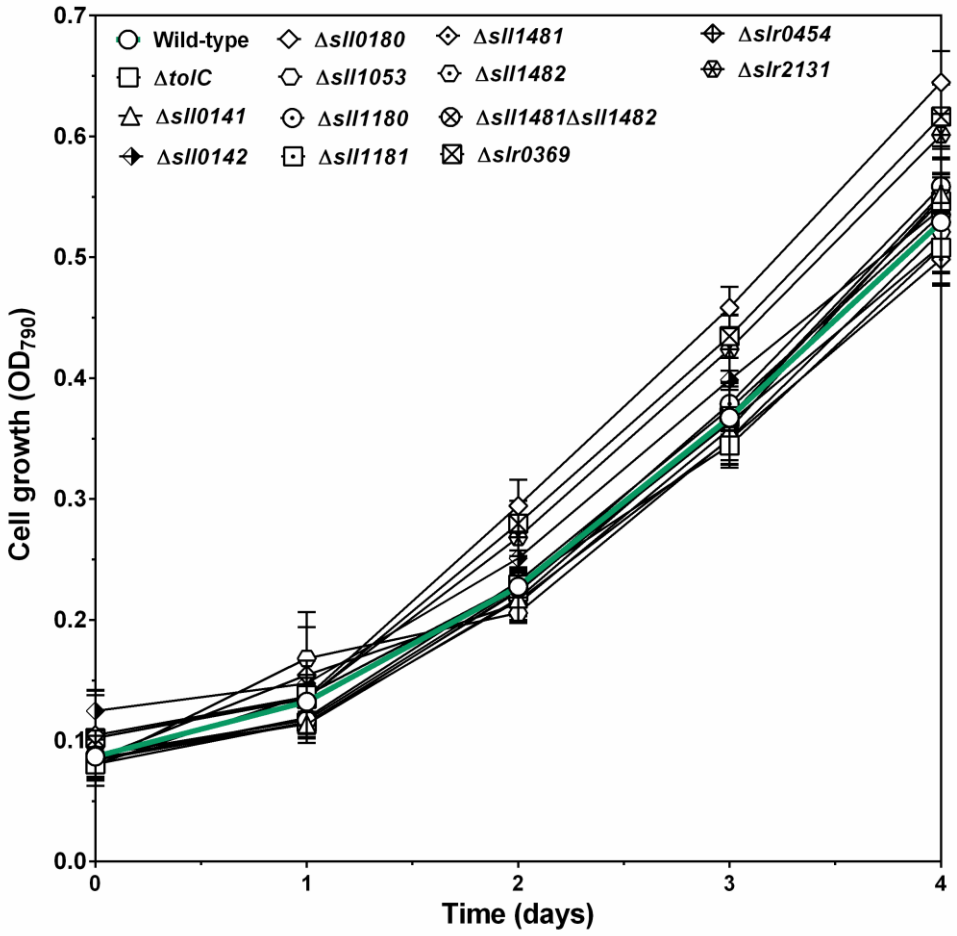


Figure 12: Growth curves of the *Synechocystis* wild-type and mutant strains.

Cyanobacterial cells were grown in 96-well microplates at 30 °C, in BG11 medium, under a 12h light (25 µmol m^{-2} s^{-1})/12h dark regimen. *Synechocystis* wild-type growth-curve is highlighted in green. For each bacterial strain, at least three independent biological replicates were performed. Error bars indicate standard deviation.

The analysis of the growth curves of all mutant strains revealed that the candidate genes do not encode essential proteins, and that cells have similar growth rates when compared to the

wild-type. Exceptionally, the single knockout mutants Δ*sll0180*, Δ*slr0369*, and Δ*slr2131* were found to grow faster than the wild-type under such conditions (**Figure 13**).

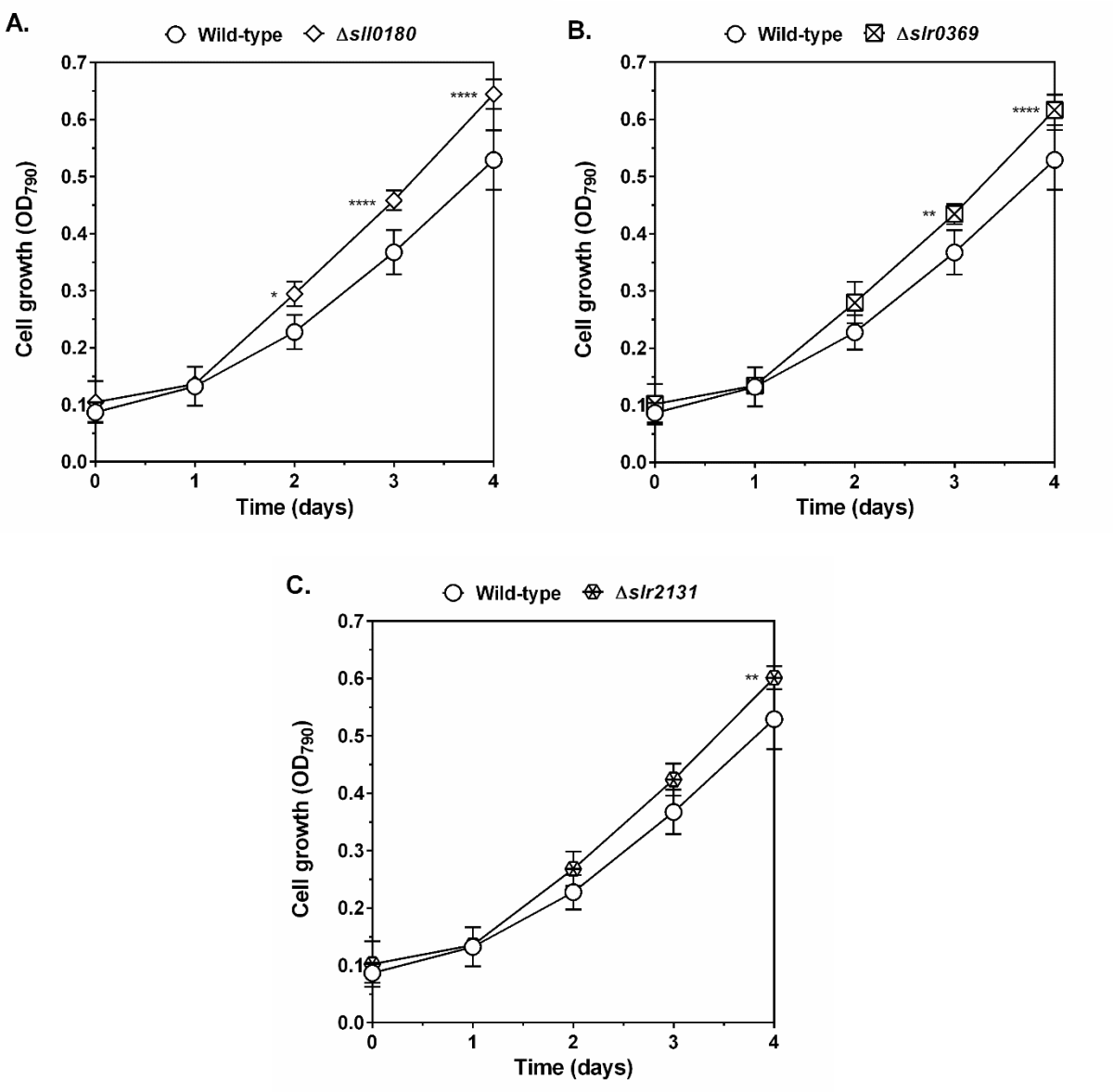


Figure 13: Growth curves of the *Synechocystis* wild-type and Δ*sll0180*, Δ*slr0369*, and Δ*slr2131* single knockout mutants.

Cyanobacterial cells were grown in 96-well microplates at 30 °C, in BG11 medium, under a 12h light (25 µmol m^{-2} s^{-1})/12h dark regimen. For each bacterial strain, at least three independent biological replicates were performed. Error bars indicate standard deviation. ns, P > 0.05; *, P ≤ 0.05; **, P ≤ 0.01; ***, P ≤ 0.001; ****, P ≤ 0.0001.

3.4 Functional characterisation of the knockout mutants

3.4.1 Evaluation of mutants' growth fitness when exposed to different exogenous compounds

RND export systems play a major part in the innate and acquired resistance of Gram-negative bacteria to several classes of structurally distinct chemical compounds that include most classes of antibiotics, as well as dyes, detergents, aromatic hydrocarbons, fatty acids, and bile salts (*32*, *95*). Seven out of the eleven candidate genes identified in this work are thought to be related to the RND superfamily of bacterial efflux pumps (**Figure 7**), and thus might have a significant role in cell survival during exposure to antimicrobial agents. For this reason, microtiter plate-based assays were performed to determine the susceptibility of the wild-type and mutant strains towards different substances characteristically secreted by RND efflux systems. A diminished growth rate in comparison to the wild-type, may suggest that the protein is involved in the export of the assayed antimicrobial compound (**Figure 14**).

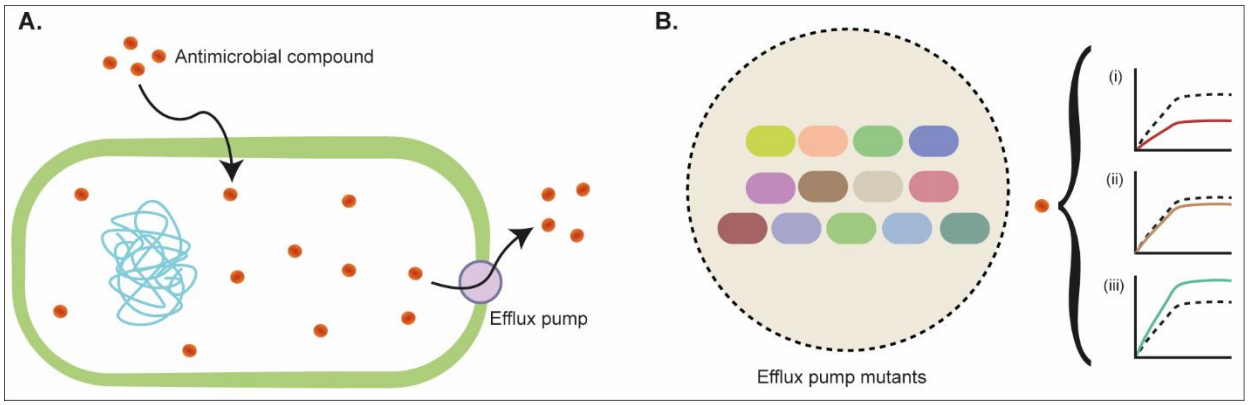


Figure 14: Screening for transporters capable of secreting exogenous compounds in *Synechocystis*.

A classical strategy to disclose the role of efflux transporters in the export of a given compound is to evaluate the susceptibility of knockout mutants towards that chemical agent. (A) If an efflux pump is engaged in the secretion of a given antimicrobial compound, it will minimise the toxic effects on cell physiology exerted by the chemical agent added exogenously. (B) Consequently, when the mutant strains are cultivated in the presence of the antimicrobial compound, three different cell growth phenotypes can be expected: (i) if a transporter is involved in compound secretion, its deletion would imply a growth inhibition in comparison with the wild-type; (ii) if not involved, there should be no growth difference between the wild-type and the mutant strain; and finally, (iii) a tolerant phenotype could result from unknown mechanisms. Adapted from Wang et *al.*, 2015 (*96*).

The concentrations of each chemical to be tested were established by testing a range of concentrations, grounded on literature searches, and on the wild-type and the Δ*tolC* mutant strain growth fitness (data not shown).

3.4.1.1 Chloramphenicol

Previously, it was determined that *Synechocystis* Δ*tolC* presents higher susceptibility to antibiotics (*60*), as it happens with other *tolC* mutants in Gram-negative bacteria (*95*). Among all the antibiotics assessed, Δ*tolC* showed the highest increase in sensitivity towards chloramphenicol; in fact, the half maximal inhibitory concentration (IC_{50}) is approximately 9 times higher in the wild-type than in the mutant strain (*60*). To explore the involvement of the candidate proteins in the secretion of chloramphenicol, cyanobacterial cells were cultivated at 30 °C and cell growth was evaluated daily, for a period of 4 days (**Figure 15**). Upon exposure to 0.05 µg mL^{-1} chloramphenicol, the wild-type exhibited the same growth pattern as in the absence of the compound, whereas the mutant Δ*tolC* displayed a growth inhibition phenotype. In order to standardize and simplify the analysis, growth inhibition was determined as a percentage, defined as the quotient between the growth rate in the presence of the exogenous compound and the growth rate in its absence (**Figure 16**).

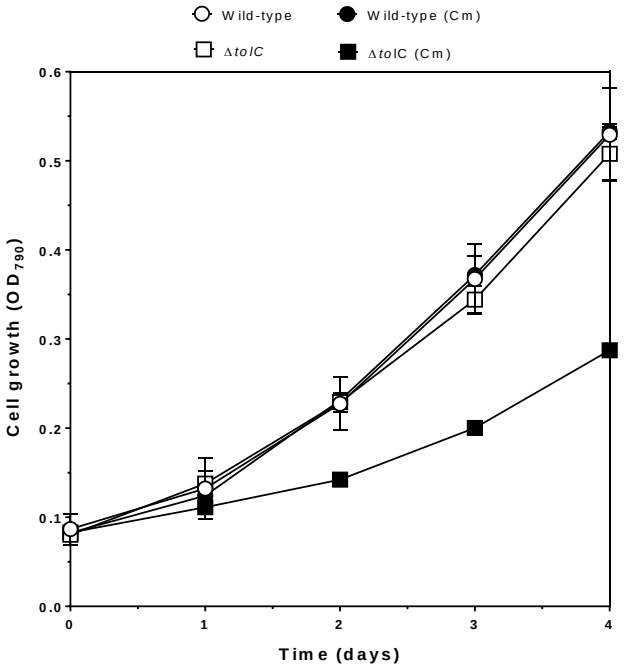


Figure 15: Growth curves of *Synechocystis* wild-type and Δ*tolC* in the presence and absence of chloramphenicol.

The wild-type and the mutant strain Δ*tolC* were either grown in BG11 medium (open symbols) or in BG11 medium supplemented with 0.05 µg mL^{-1} chloramphenicol (full symbols). For each bacterial strain, at least three independent biological replicates were carried out. Error bars indicate standard deviation.

From an ecological standpoint, it has been widely accepted that antibiotics are produced by many bacterial strains with the aim of inhibiting the growth of microbial competitors. In contrast, one might assume that antibiotic resistance determinants, such as efflux pumps, should serve to

limit the activity of antibiotics (*40*). Although this may happen, an alternative theory has been recently proposed and explores the hypothesis that many organic compounds synthesised and secreted by microorganisms (e.g. antibiotics, cationic peptides, and other low-molecular-weight metabolites) could function as cell-signalling molecules, shaping the structure of microbial communities (*97*, *98*). According to this hypothesis, antimicrobial compounds will have a hormetic effect, beneficial at the sub-inhibitory concentrations found in most natural ecosystems and harmful at the high concentrations used for therapeutic purposes (*98*, *99*). For example, when *E. coli* strain MG1655 was exposed to tetracycline concentrations much lower than the minimum inhibitory concentration (MIC) a hormetic increase was observed: the colony forming unit (CFU) number rose up to 141% and 121% as compared to the control (*100*). In the same way, it is possible to observe that when exposed to 0.05 µg mL^{-1} chloramphenicol (**Figure 16A**), *Synechocystis* wild-type cells display a hormetic response and present an increased fitness when compared to cells grown in BG11 medium; hence, the negative growth inhibition percentage. It should be noted that the IC_{50} of chloramphenicol in wild-type *Synechocystis* has been determined to be 1.8±0.5 µg mL^{-1} (*60*). Similar phenotypes were observed when wild-type cells were exposed to the lowest concentration of SDS and Triton X-100 assayed (see below, **Figure 17A** and **Figure 18A**). Despite the synthetic nature of these detergents, their action on cell physiology is similar to that of many bacterial surfactants (*101*), presumably eliciting similar response mechanisms. Regarding the knockout mutants, Δ*tolc* proved to be hypersensitive to all concentrations tested, whereas the remaining mutants exhibited various phenotypes, according to the relevance of the insertionally inactivated gene in efflux-mediated resistance to the assayed compound.

There were no significant differences revealed in the sensitivity of the mutants carrying disrupted alleles of genes putatively related to the T1SS (Δ*sll1180* and Δ*sll1181*) or homologous to the heterocyst development genes from *Anabaena* sp. PCC 7120 (Δ*sll1481* and Δ*sll1482*), regardless of the assayed chloramphenicol concentration (**Figure 16**). Likewise, the same was observed for the knockout mutants deficient in the putative RND IM transporter Slr0369 (Δ*slr0369*), and in the putative RND adaptor proteins Sll0141 and Sll0180 (Δ*sll0141* and Δ*sll0180*). In contrast, the mutants Δ*sll1053* and Δ*slr2131* presented a growth impairment similar to Δ*tolC*. Curiously, Δ*slr0454* is approximately 2 times more susceptible than the wild-type when cultivated in the presence of 5 times more chloramphenicol (0.25 µg mL^{-1}) (**Figure 16B**), even though it was among the least susceptible mutants when exposed to a lower concentration (**Figure 16A**).

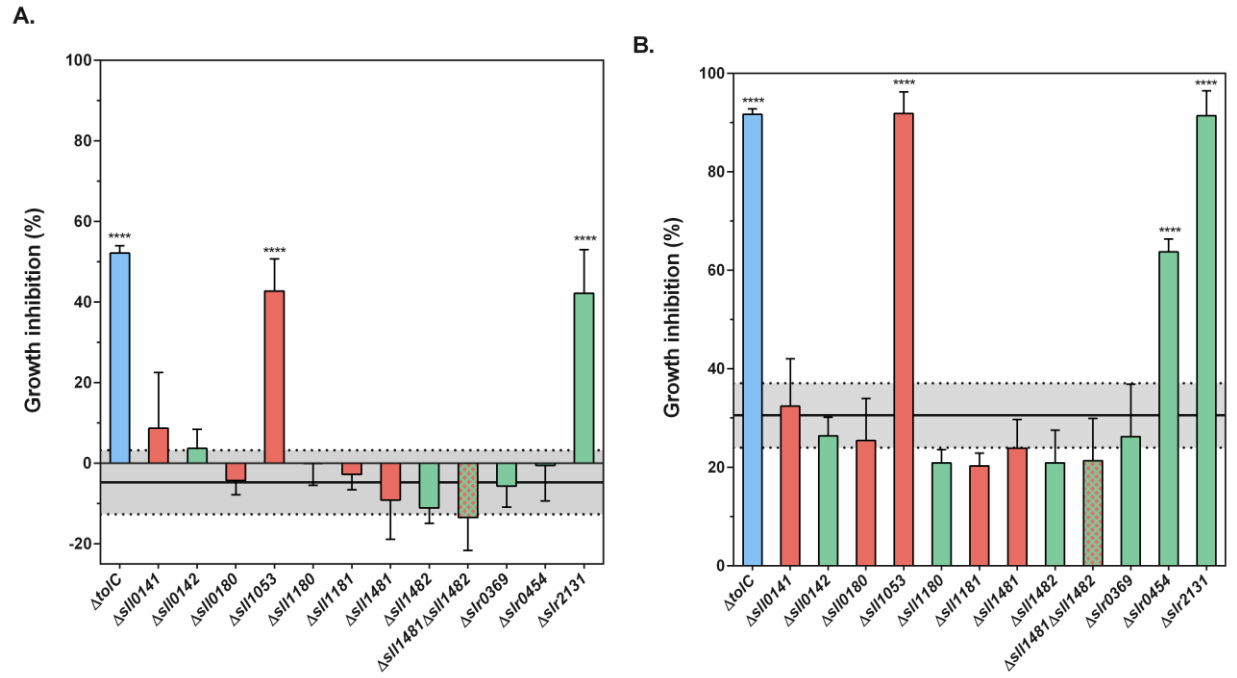


Figure 16: Growth inhibition analysis of *Synechocystis* wild-type and mutant strains when exposed to exogenously added chloramphenicol.

All bacterial strains were grown at 30 °C for 4 days in BG11 medium with (A) 0.05 µg mL^{-1} or (B) 0.25 µg mL^{-1} chloramphenicol. The shadowed area represents the growth inhibition of the wild-type: the mean value is depicted by a full line and the upper and lower standard deviation limits by dashed lines. Mutants lacking functional outer membrane proteins (OMPs), periplasmic adaptor proteins (PAPs) or inner membrane components (IMCs) are depicted in blue, red and green, respectively. A mutant lacking both the PAP and the IMC is depicted in red and green. Error bars represent the standard deviations of at least three independent biological replicates. ns, P > 0.05; *, P ≤ 0.05; **, P ≤ 0.01; ***, P ≤ 0.001; ****, P ≤ 0.0001.

In light of the obtained results, it is possible to suggest that the IM transporter encoded by *slr2131* and the PAP encoded by *sll1053* form the core RND translocase complex responsible for the secretion of chloramphenicol. Nonetheless, the multi-resistance potential provided by RND efflux pumps is typically enhanced by the presence of multiple complexes in a single bacterium that sometimes have overlapping substrate specificities and can act sequentially when others fail (*36*, *95*). Accordingly, the growth phenotype exhibited by Δ*slr0454* upon exposure to a higher concentration of chloramphenicol, seems to suggest that similar mechanisms also exist in *Synechocystis*. Thus, Slr0454 seems to function as an alternative and complementary efflux route that is used when the capacity of the main efflux pump to secrete chloramphenicol is exceeded.

3.4.1.2 Sodium dodecyl sulphate (SDS)

SDS is an anionic surfactant that disrupts biological membranes and denatures proteins by breaking protein:protein interactions. This denaturing detergent is also a known substrate of TolC-

dependent efflux systems in other Gram-negative bacteria (*95*). For the lower concentration of SDS assayed, only Δ*tolC*, Δ*sll0142,* Δ*sll0180* and Δ*slr0369* presented significantly increased susceptibility in comparison to the wild-type (**Figure 17A**). However, growth of Δ*tolC* is already completely inhibited at such low concentration, suggesting that TolC-dependent efflux systems are of essential importance to counteract the detrimental effects exerted by this detergent on cell physiology and viability. On the other hand, Δ*sll0142*, Δ*sll0180* and Δ*slr0369* are about 5 times less susceptible than Δ*tolC*, which indicates that other mechanisms of resistance to SDS might exist in this cyanobacterium, namely additional TolC-dependent export systems. Nonetheless, it is possible to suggest that the RND IM transporters encoded by *sll0142* and *slr0369* are involved in the secretion of SDS, alongside with the RND adaptor protein encoded by *sll0180*.

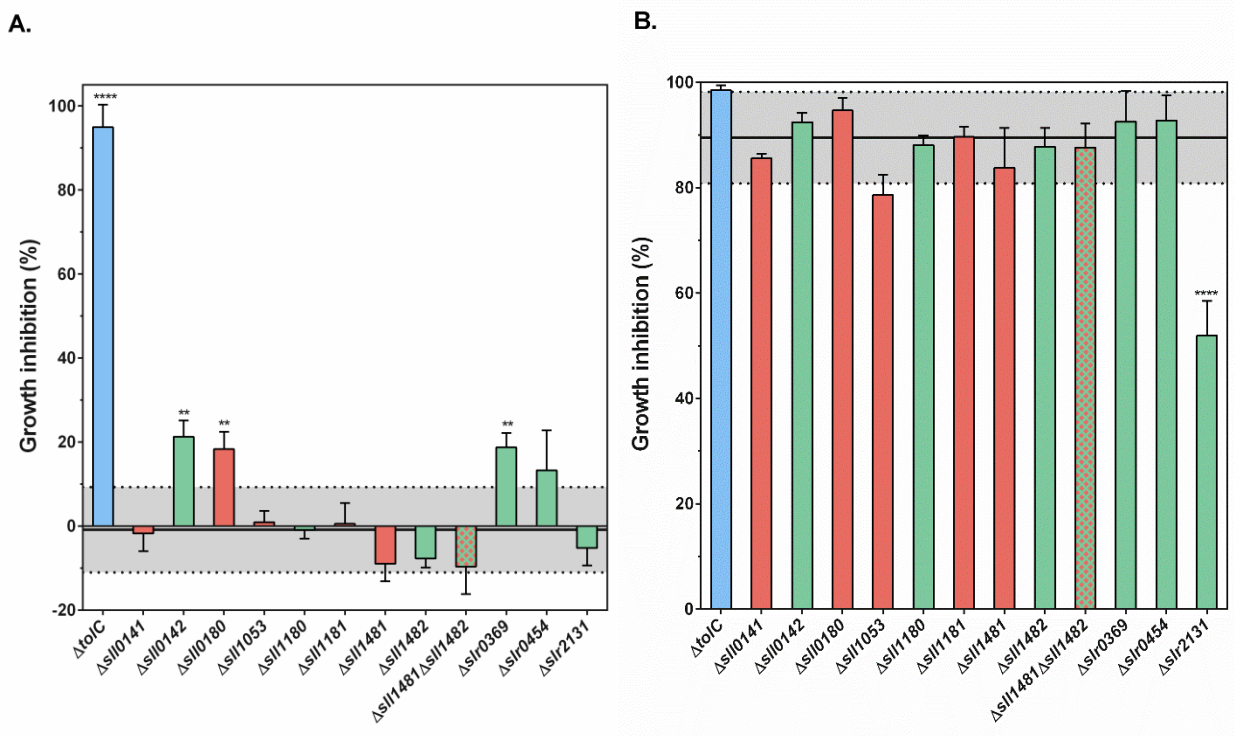


Figure 17: Growth inhibition analysis of *Synechocystis* wild-type and mutant strains when exposed to exogenously added sodium dodecyl sulphate (SDS).

All bacterial strains were grown at 30 °C for 4 days in BG11 medium with (A) 0.0025 % (w/v) or (B) 0.0250 % (w/v) SDS. The shadowed area represents the growth inhibition of the wild-type: the mean value is depicted by a full line and the upper and lower standard deviation limits by dashed lines. Mutants lacking functional outer membrane proteins (OMPs), periplasmic adaptor proteins (PAPs) or inner membrane components (IMCs) are depicted in blue, red and green, respectively. A mutant lacking both the PAP and the IMC is depicted in red and green. Error bars represent the standard deviations of at least three independent biological replicates. ns, P > 0.05; *, P ≤ 0.05; **, P ≤ 0.01; ***, P ≤ 0.001; ****, P ≤ 0.0001.

Remarkably, the single knockout mutant Δ*slr2131* proved to be more tolerant than the wild-type or any of the other mutants upon exposure to 0.0250 % (w/v) SDS (**Figure 17B**). TolC is a

promiscuous OMP that exhibits affinity to various translocase complexes and is able to interact with them in a very dynamic process (*102*). Therefore, the different IM translocases complexes have to compete with each other for recruitment of TolC, either by being constitutively expressed in larger quantities or by displaying higher affinity towards TolC (*102*). In such scenario, one might consider that dominant transporters occupy most of a cell's TolC pool in the wild-type, whereas TolC molecules become more available for non-dominant transporters when dominant transporters are mutated (**Figure 18**). In the single knockout mutant Δ*slr2131*, it is therefore possible that the main transporters engaged in secretion of SDS are able to recruit TolC faster or better than in the wild-type background, which improves tolerance to this anionic detergent.

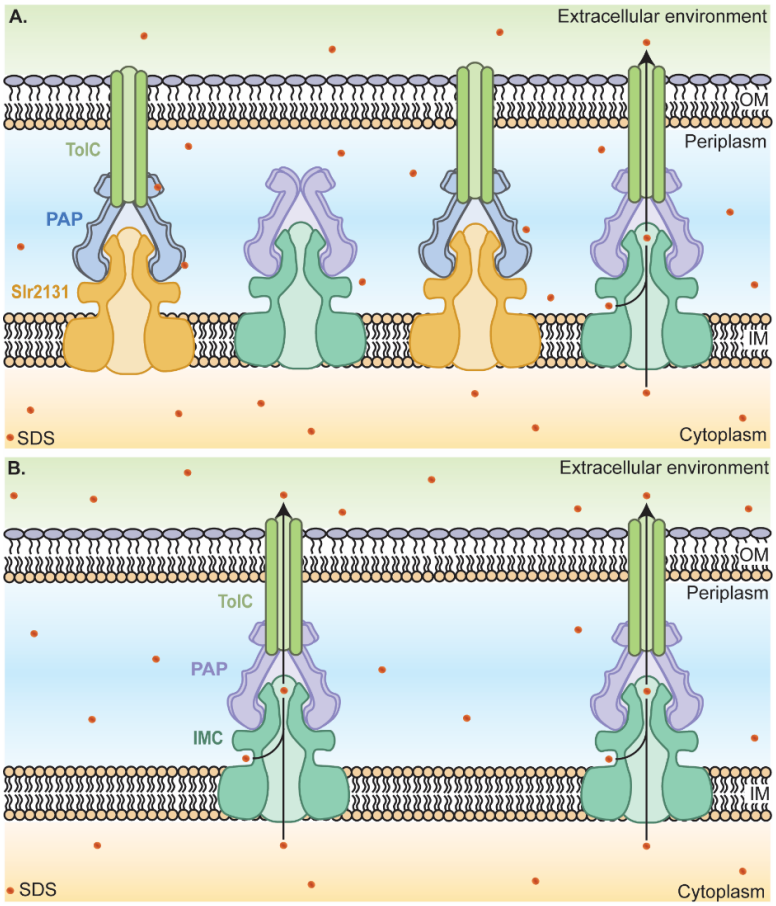


Figure 18: Proposed mechanism for sodium dodecyl sulphate (SDS) secretion in *Synechocystis*.

(A) In the wild-type strain, the different inner membrane (IM) translocases complexes, composed by an IM component and a periplasmic adaptor protein (PAP), have to compete with each other for recruitment of TolC. Dominant transporters, such as slr2131 (yellow and blue), occupy most of a cell's TolC pool in the wild-type, forming a tripartite complex with the outer membrane (OM) channel TolC. However, this multiprotein complex does not act on SDS secretion and limits the capacity of other transporters (green and purple) to interact with TolC and expel this antimicrobial compound. (B) In the single knockout mutant Δ*slr2131*, the deletion of this major transporter increases the availability of TolC for SDS transporters to form tripartite complexes with TolC for direct efflux of SDS. Adapted from Wang et *al.*, 2015 (*96*).

3.4.1.3 Triton X-100

Non-denaturing detergents such as Triton X-100 generally do not disrupt native interactions and structures of water-soluble proteins. The time courses of cell growth confirmed that disruption of *tolC* resulted in severe growth inhibition when Triton X-100 was present (**Figure 19**), once more emphasising the importance of this OMP as an effective strategy to overcome the toxicity induced by antimicrobial compounds. As with the other surfactant assayed in this work, none of the mutants' growth inhibition approximated that shown by Δ*tolC*, which suggests that other mechanisms of resistance to Triton X-100 might exist in *Synechocystis*, namely additional TolC-dependent export systems. Nonetheless, the growth of the mutant strains Δ*sll0180* and Δ*slr0369* was significantly inhibited in comparison to the wild-type (**Figure 19A**). Moreover, the mutants deficient in Sll0142, Sll1053, and Slr0454 were also more susceptible than the wild-type when exposed to the highest concentration assayed (**Figure 19B**).

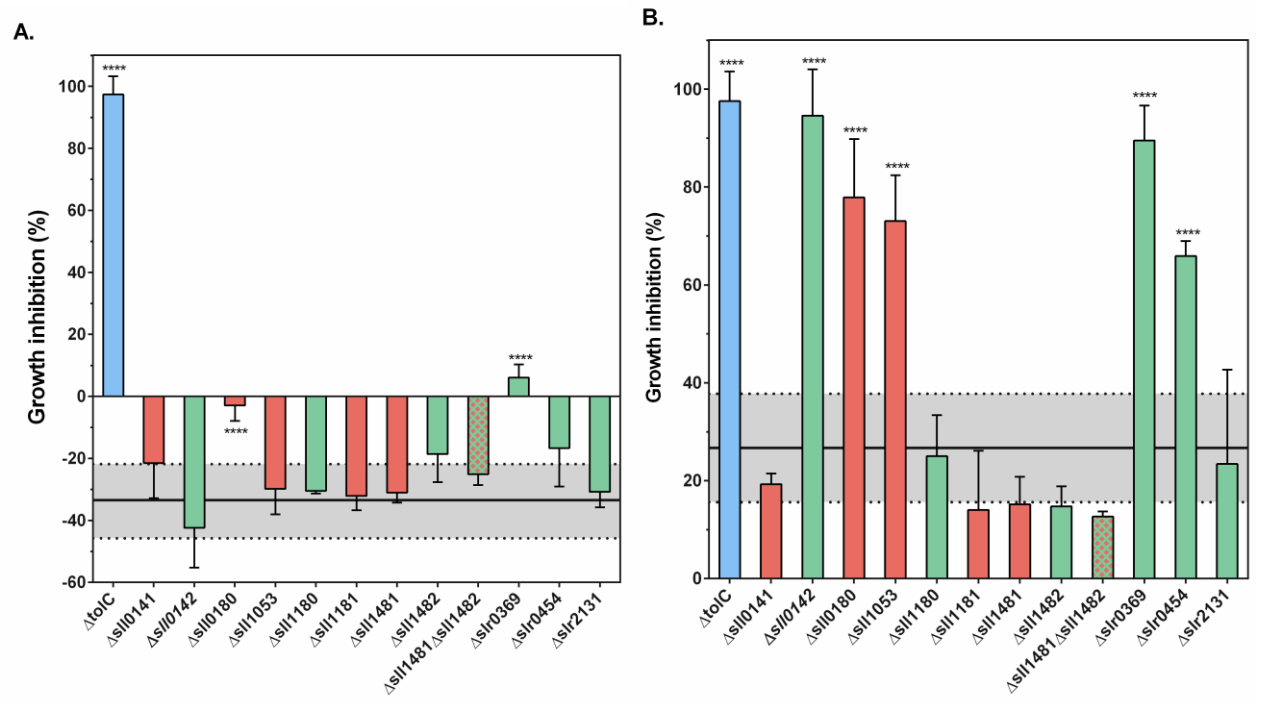


Figure 19: Growth inhibition analysis of *Synechocystis* wild-type and mutant strains when exposed to exogenously added Triton X-100.

All bacterial strains were grown at 30 °C for 4 days in BG11 medium with (A) 0.125 % (v/v) or (B) 1.250 % (v/v) Triton X-100. The shadowed area represents the growth inhibition of the wild-type: the mean value is depicted by a full line and the upper and lower standard deviation limits by dashed lines. Mutants lacking functional outer membrane proteins (OMPs), periplasmic adaptor proteins (PAPs) or inner membrane components (IMCs) are depicted in blue, red and green, respectively. A mutant lacking both the PAP and the IMC is depicted in red and green. Error bars represent the standard deviations of at least three independent biological replicates. ns, P > 0.05; *, P ≤ 0.05; **, P ≤ 0.01; ***, P ≤ 0.001; ****, P ≤ 0.0001.

Based on the obtained results, it is possible to suggest that the PAP Sll0180 and the IM transporter Slr0369 may form a tripartite assembly with TolC, mediating the secretion of Triton X-100 to the extracellular medium. As formerly discussed, IM translocase complexes might have overlapping substrates specificities and interact with different protein components, modulating substrate specificity and recognition (*36*, *38*, *95*). Herein, it has been proposed that the Sll1053 is involved in the secretion of chloramphenicol by interacting with the IM transporter Slr2131. Moreover, Δ*sll1053* exhibits a growth inhibition phenotype similar to that presented by the Δ*sll0142* and Δ*slr0454* knockout mutant when exposed to 1.250 % (v/v) of Triton X-100. These observations suggest this PAP may mediate secretion of Triton X-100 either by interacting with Sll0142 and/or Slr0454, or by coupling with an additional and yet unidentified IMC.

3.4.2 Efflux activity of the mutants in the presence of ethidium bromide

An agar-based screening assay was carried out with the aim of characterising the efflux behaviour of the knockout mutants in the presence of the fluorescent dye ethidium bromide (**Figure 20**), which is a known substrate of RND efflux pumps in Gram-negative bacteria (*95*).

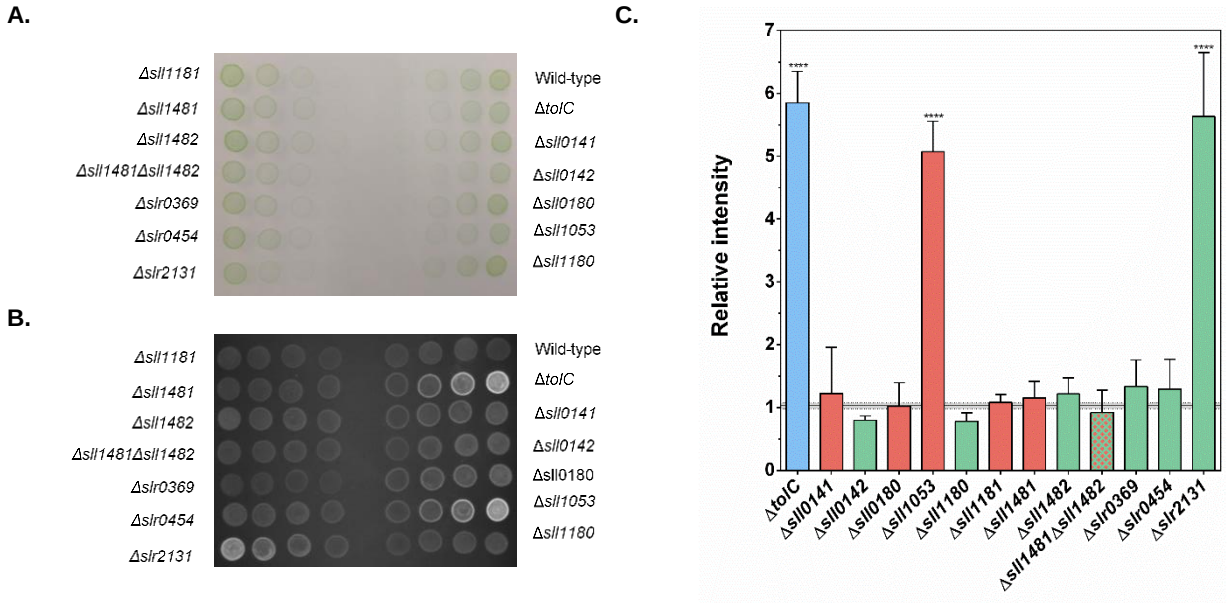


Figure 20: Evaluation of the efflux activity of *Synechocystis* wild-type and mutant strains in the presence of ethidium bromide. Efflux activity was assayed by an ethidium bromide agar-based screening methodology. (A) A series dilution of *Synechocystis* cells was pipetted onto solid BG11 medium supplemented with 0.5 µg mL^{-1} ethidium bromide. (B) The cells were incubated for 24 hours, at 25 °C, under a 16h light (25 µmol photons m-2 s-1))/8h dark regimen, before the plates were observed under ultraviolet (UV) light; images were acquired using a GelDoc documentation system (Bio-Rad). (C) Fluorescence intensity was later quantified using the ImageJ software. Mutants lacking functional outer membrane proteins (OMPs), periplasmic adaptor proteins (PAPs) and inner membrane components (IMCs) are depicted in blue, red and green, respectively. Mutants lacking PAPs and IMCs are depicted in red and green.

Upon entrance into the cell, ethidium bromide will intercalate between the DNA base pairs and give rise to fluorescent signal when cells are exposed to ultraviolet (UV) light. Nevertheless, Gram-negative cells harbour secretion mechanisms that mediate the efflux of the compound, and so the higher the fluorescence signal, the less efficient the cell is in secreting ethidium bromide. When the wild-type and the Δ*tolC* strains were grown in solid BG11 medium supplemented with ethidium bromide, it was observed that mutant cells presented a higher fluorescence signal in comparison to wild-type cells (**Figure 20**). This observation suggests that there is an impairment of ethidium bromide transport functions in the Δ*tolC* mutant, similarly to what has been reported earlier (*103*, *104*). The role of the putative transporters herein identified was also investigated and it was possible to observe fluorescence levels in the Δ*sll1053* and Δ*slr2131* knockout mutants similar to those quantified for Δ*tolC*. On the other hand, the remaining mutants exhibited fluorescence levels comparable to the wild-type. Therefore, it seems the presumptive candidate proteins encoded by the genes *sll1053* and *slr2131* play a role in the efflux of ethidium bromide, conceivably forming a tripartite assembly with TolC.

3.4.3 Protein secretion

Previously, it was shown that a TolC-like protein (Slr1270) mediates protein secretion in *Synechocystis* as the overall composition of the exoproteome of a knockout mutant (Δ*tolC*) is significantly different from the wild-type (*60*). In the exoproteome of this mutant strain several Coomassie blue stained bands are missing, including the band corresponding to Sll1951 (**Figure 21**, band highlighted with an asterisk), the main constituent of the S-layer (*60*,*105*), indicating the secretion of this protein occurs by means of a TolC-dependent mechanism. Herein, two putative components of the T1SS were bioinformatically identified in *Synechocystis*: Sll1180 and Sll1181. Their role and that of the remaining candidate proteins in the secretion of proteins was studied by cultivating cells in photoautotrophic conditions and examining the extracellular protein content present in the growth medium of the various strains (**Figure 21**). In the exoproteome of the knockout mutants Δ*sll0141*, Δ*sll0180*, Δ*sll1180* and Δ*sll1181* several Coomassie blue stained bands are missing, including the band corresponding to Sll1951, which suggests *Synechocystis* has to retain function of these proteins in order to accomplish protein transport in a TolC-dependent manner. Recently, it has been suggested that the TriABC-OpmH triclosan extrusion system of *P. aeruginosa* requires two different PAPs for functional activity, which is most easily explained with a 3:6:3 stoichiometry where the two different PAPs occupy two non-equivalent surface grooves on OpmH (*108*, *109*). Similarly, the ZrpADBC pump of *Serratia* sp. ATCC 39006 possesses two PAPs (ZrpA and ZrpD), which may assemble with a 3:6:3 stoichiometry (*110*).

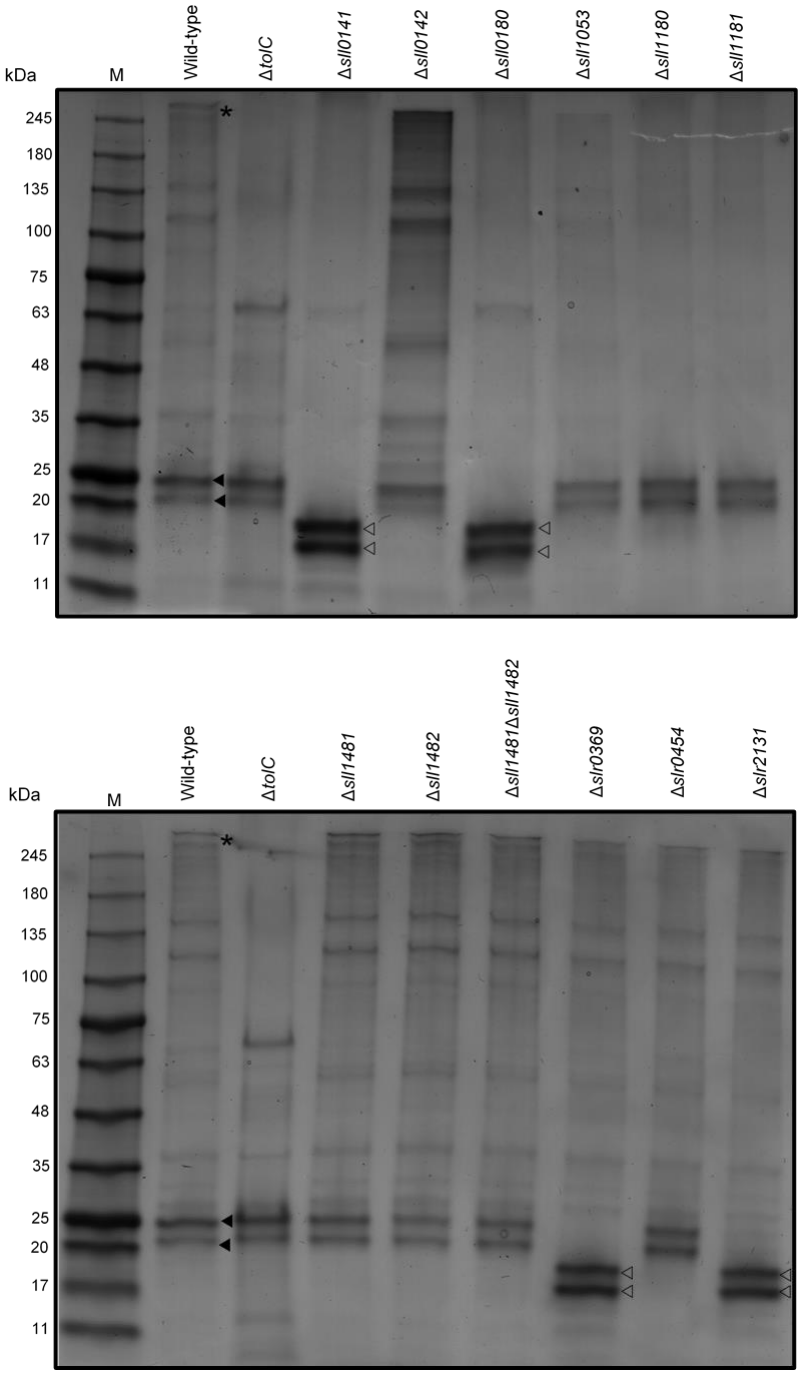


Figure 21: Exoproteome analysis of *Synechocystis* wild-type and mutant strains.

Coomassie blue stained 4-15% (w//v) polyacrylamide gels showing the proteins accumulating in the extracellular environment of *Synechocystis* wild-type and mutant strains. The band highlighted with an asterisk (*) in the exoproteome of the wild-type strains was previously identified as the protein Sll1951, the main constituent of the S-layer of *Synechocystis* (*60,105*). Bands indicated by full arrowheads (◀) in the exoproteome of Δ*tolC* were formerly identified as the pili-forming protein, PilA1 (Sll1694) (*60*). Bands indicated by open arrowheads (◁) were excised from the gel for further analysis and protein identification. The pattern of bands displayed in the gels was consistent across two independent biological replicates. M, NZYColour Protein Marker II. Molecular weights are indicated on the left (kDa).

In addition, the purified soluble AcrA and MexA proteins can also form oligomers under certain conditions (*111*). Even though the stoichiometry and oligomeric state of the PAP assembled with the remaining partners of the multiprotein complex remains unknown, the propensity to oligomerisation *in vitro*, cross-linked dimers and trimers *in vivo* and the presence of paired PAPs in bacterial genomes all support the notion that in transport complexes PAPs function as homo- or hetero-oligomers (*42*). Analysis of the exoproteome of the Δ*tolC*, Δ*sll0141*, Δ*sll0180*, Δ*sll1180* and Δ*sll1181* knockout mutants indicates that the proteins encoded by the insertionally inactivated genes are involved in the secretion of proteins, namely substrates typical of the T1SS (e.g. Sll1951). Thus, it is possible to suggest that Sll1180, the putative ABC-transporter, might interact with three adaptor proteins, Sll0141, Sll0180 and Sll1181, to mediate the transport of proteins across the cell envelope in a TolC-dependent mechanism. Alternatively, the absence of S-layer in the Δ*sll0141* and Δ*sll0180* mutant strains may also be the result of Sl0141 and Sll0180 function(s) in molecular mechanisms that ensure proper assembly of the Sll1180-Sll1181-TolC complex, and not on the direct mechanism of protein secretion. Nonetheless, additional experimental assays will have to be carried out with the aim of testing these hypotheses.

In the extracellular environment of the wild-type and the Δ*tolC* mutant it is also possible to observe two bands located between 20 and 25 kDa (**Figure 21**, bands indicated with full arrowheads), which had been formerly identified as pilin A1 (PilA1, Sll1694), the main structural component of pili in *Synechocystis* (*60*, *112*). Interestingly, these two bands are absent in the exoproteome of the single knockout mutants Δ*sll0141*, Δ*sll0180*, Δ*slr0369* and Δ*slr2131.* In alternative, two bands of lower molecular weight, located between 20 and 17 kDa (**Figure 21**, bands indicated with open arrowheads), can be found. The exoproteins of lower molecular weight had never been identified in the extracellular environment of the wild-type or the mutant Δ*tolC*; therefore, the bands of interest were excised from the gel, in-gel trypsin digested and the peptides present therein were identified by mass spectrometry. Remarkably, the bands indicated by open arrowheads were also identified as PilA1, suggesting that when separated in SDS-polyacrylamide gels, Sll1694 presents different electrophoretic mobility depending on the genetic background of the analysed strain. Since SDS-polyacrylamide gel electrophoresis separates proteins only by their size and the *sll1694* gene is predicted to encode a 17.58 kDa protein, analysis of the pattern of bands and their estimated molecular weight in comparison with the molecular weight marker, suggests the existence of variants of Sll1694 caused by post-translational modifications. As a matter of fact, post-translational modifications of pilin subunits are thought to be critical for the biogenesis of functional pili in *Synechocystis* (*112*), and include O-linked glycosylation of pilin

subunits and trimethylation of pilin at the C-terminal lysine (*113–115*). Moreover, PilA1 is produced as precursor protein that has its N-terminal leader sequence cleaved by the prepilin peptidase, PilD (*112*). Thus, in addition to the analysis of the exoproteome (**Figure 21**), the same samples were electrophoresed on similar gels and stained using the periodic acid-Schiff (PAS) method (**Figure 22**), which allows to specifically stain glycosylated proteins. The bands located between 20 and 25 kDa are both glycosylated, whereas no bands could be observed in the lanes corresponding to the concentrated exoproteome samples of the mutants Δ*sll0141*, Δ*sll0180*, Δ*slr0369* and Δ*slr2131*, confirming PilA1 is not glycosylated in these mutant *Synechocystis* strains.

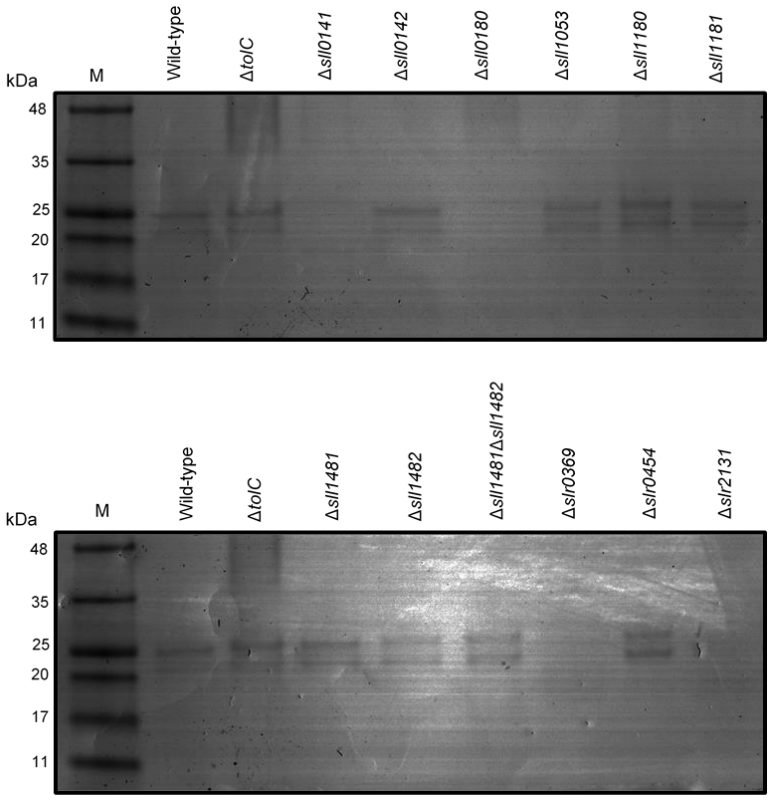


Figure 22: Analysis of glycosylated proteins by SDS-polyacrylamide gel electrophoresis.

Concentrated exoproteome samples were separated on 4-15% (w/v) gradient polyacrylamide gels and stained using the Glycoprotein Staining Kit (Pierce). M, NZYColour Protein Marker II (NZYTech). Molecular weights are indicated on the left.

Therefore, the experimental evidences gathered suggest that in the exoproteome of the wild-type only glycosylated variants of prepilin (molecular weight of approximately 25 kDa) and mature PilA1 (molecular weight of approximately 20 kDa) can be found, in agreement to what has been described earlier (*113*). Furthermore, mutants Δ*tolC*, Δ*sll0142*, Δ*sll1053*, Δ*sll1180*, Δ*sll1181*, Δ*sll1481*, Δ*sll1482*, Δ*sll1481* Δ*sll1482* and Δ*slr0454* present similar pilin variants. However, in the extracellular medium of the single knockout mutants Δ*sll0141*, Δ*sll0180*, Δ*slr0369* and Δ*slr2131*

only non-glycosylated variants of prepilin (molecular weight of approximately 20 kDa) and mature PilA1 (molecular weight of approximately 17 kDa) can be observed. Moreover, the non-glycosylated variants of Sll1694 seem to accumulate in higher amounts than in other analysed strains.

Disruption of the gene *slr1443*, which encodes a protein homologous to eukaryotic serine/threonine protein kinases, has been found to be involved in post-translational modifications of PilA1 (*113*). In the absence of this protein, immunoblots of whole cells lysates revealed the presence of two variants of PilA1, organised in a pattern similar to what has been herein described for the Δ*sll0141*, Δ*sll0180*, Δ*slr0369*, and Δ*slr2131* single knockout mutants, suggesting these proteins are somehow involved in the glycosylation of PilA1. The role played by the putative PAPs Sll0141 and Sll0180, and presumptive IMCs Slr0369 and Slr2131 on the glycosylation of pilin in *Synechocystis*, and their relation to the function of Slr1443 is unclear. Additional work needs to be performed in order to clarify the intricate mechanisms by which multidrug efflux, T1SS-dependent protein secretion, and pilin maturation and export are related. Nonetheless, the experimental evidences provided here support the possibility of TolC-independent functional roles among IM translocase components that also seem to be involved in TolC-dependent secretion mechanisms. Such extensive functional diversity was not yet reported in other Gram-negative microorganisms.

In addition, when analysing total protein extracts of *Synechocystis* wild-type and mutant strains separated by electrophoresis on SDS-polyacrylamide gels (**Figure 23**), it is possible to observe that no major differences were detected between the two extracts. Nevertheless, two differences were noteworthy: (i) the band corresponding to the Sll1951 protein, which presents a molecular weight higher than 250 kDa in wild-type extracts, is absent in the Δ*tolC*, Δ*sll0141*, Δ*sll0180*, Δ*sll1180* and Δ*sll1181* mutants (**Figure 23**, band highlighted by a star), in agreement with the observations previously described (**Figure 21**); (ii) a band of approximately 100 kDa, formerly identified as the phycobilisome core-membrane linker polypeptide ApcE (Slr0335) (*60*), presents different electrophoretic mobility in the Δ*tolC*, Δ*sll0141*, Δ*sll0180*, Δ*sll1180* and Δ*sll1181* mutants when compared to the wild-type (**Figure 23**, band highlighted by a arrowhead). It has been proposed that ApcE acts both as a terminal energy acceptor, assuring energy transferences from the phycobilisomes to the chlorophyll *a* associated with the photosystems, and as a linker polypeptide that stabilises the phycobilisome core architecture (*106*). A mutant lacking functional Sll1951, Δ*sll1951*, was shown to accumulate more UV-absorbing metabolites in its extracellular environment than in the wild-type (*105*). Moreover, it has been demonstrated that S-layer constituents accumulate at high proton fluxes in *Prochlorothrix hollandica* (*107*), suggesting this

outermost coating is involved in the protection of cells against UV light. Thus, one might hypothesize that in the absence of this layer, *Synechocystis* cells are more exposed to photo-oxidative stress and modifications of the ApcE protein might help dissipate the excess of photons.

As some of the mutant strains present significant differences in the exoproteome, the biochemical analyses presented above were complemented with observations of the cell surface structures of selected strains by negative staining TEM. In wild-type cells, it was possible to observe a reticulated lattice as the outermost cell wall structure in wild-type cells, consistent with the formerly described S-layer (*105*), whereas the Δ*sll0141*, Δ*sll0180,* Δ*sll1180*, and Δ*sll1181* single knockout mutants lack such structure (**Figure 24**), similarly to what has been reported earlier for the Δ*tolC* mutant (*60*). Interestingly, numerous pili can be found on the surface of the Δ*sll0141*, Δ*sll0180*, Δ*slr0369*, and Δ*slr2131* mutant cells, sometimes forming a capsule-like amorphous mass that encircles the cell (**Figure 24**, highlighted by a star). On the other hand, wild-type cells seem to possess less pili in comparison with these mutants, which is also supported by the higher accumulation of prepilin and PilA1 in the extracellular medium of these knockout mutants (**Figure 21**). These results seem to suggest that glycosylation of PilA1 is essential for proper pili assembly and structure. However, it remains to be determined whether the function of these altered pili is retained.

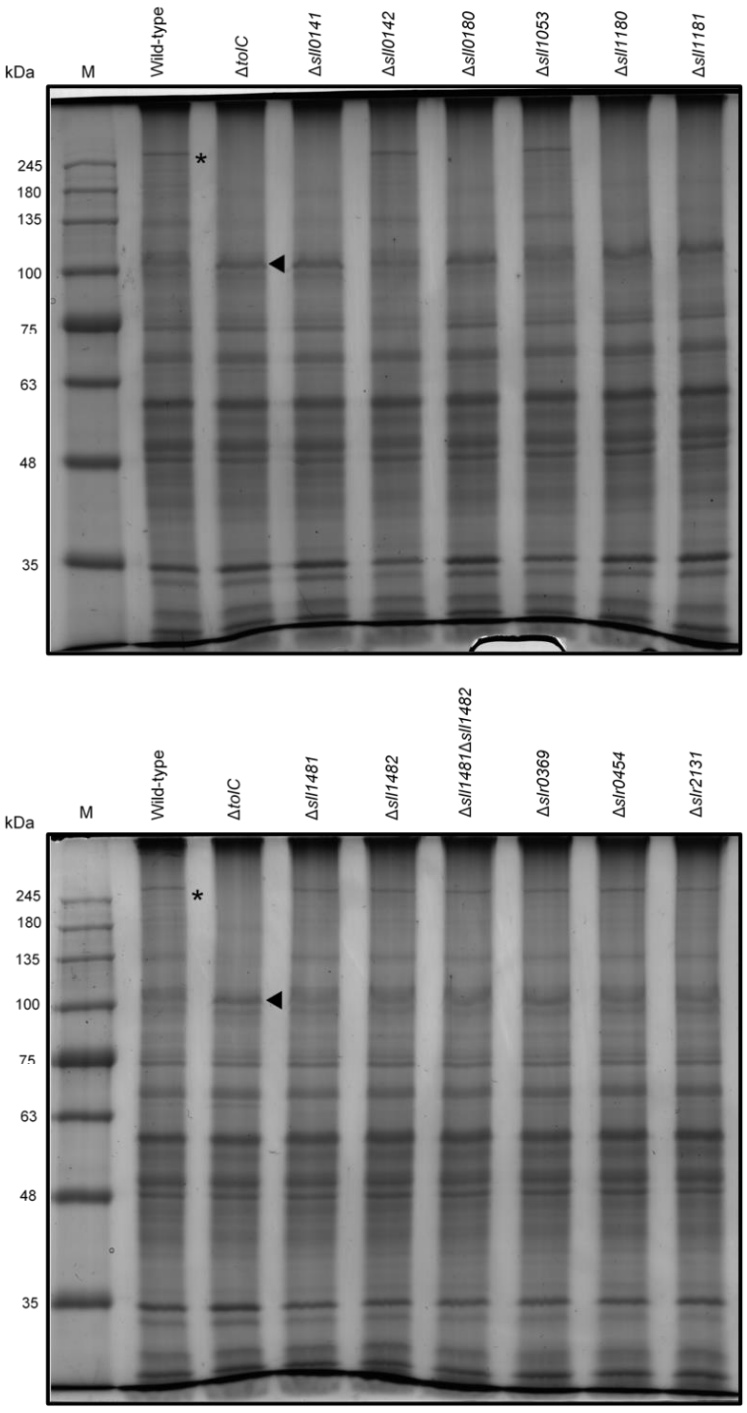


Figure 23: Total protein extracts of *Synechocystis* wild-type and mutant strains.

Coomassie blue stained 8% (w/v) SDS-polyacrylamide gels showing total protein extracts of *Synechocystis* wild-type and mutant strains. The band highlighted with an asterisk (*) was previously identified as the protein Sll1951, the main constituent of the S-layer of *Synechocystis* (*60*). Bands indicated by full arrowheads (◀) were formerly identified by mass spectrometry as ApcE (Slr0335) (*60*). The pattern of bands displayed in the gels was consistent across two independent biological replicates. M, NZYColour Protein Marker II. Molecular weights are indicated on the left (kDa).

A. Wild-type

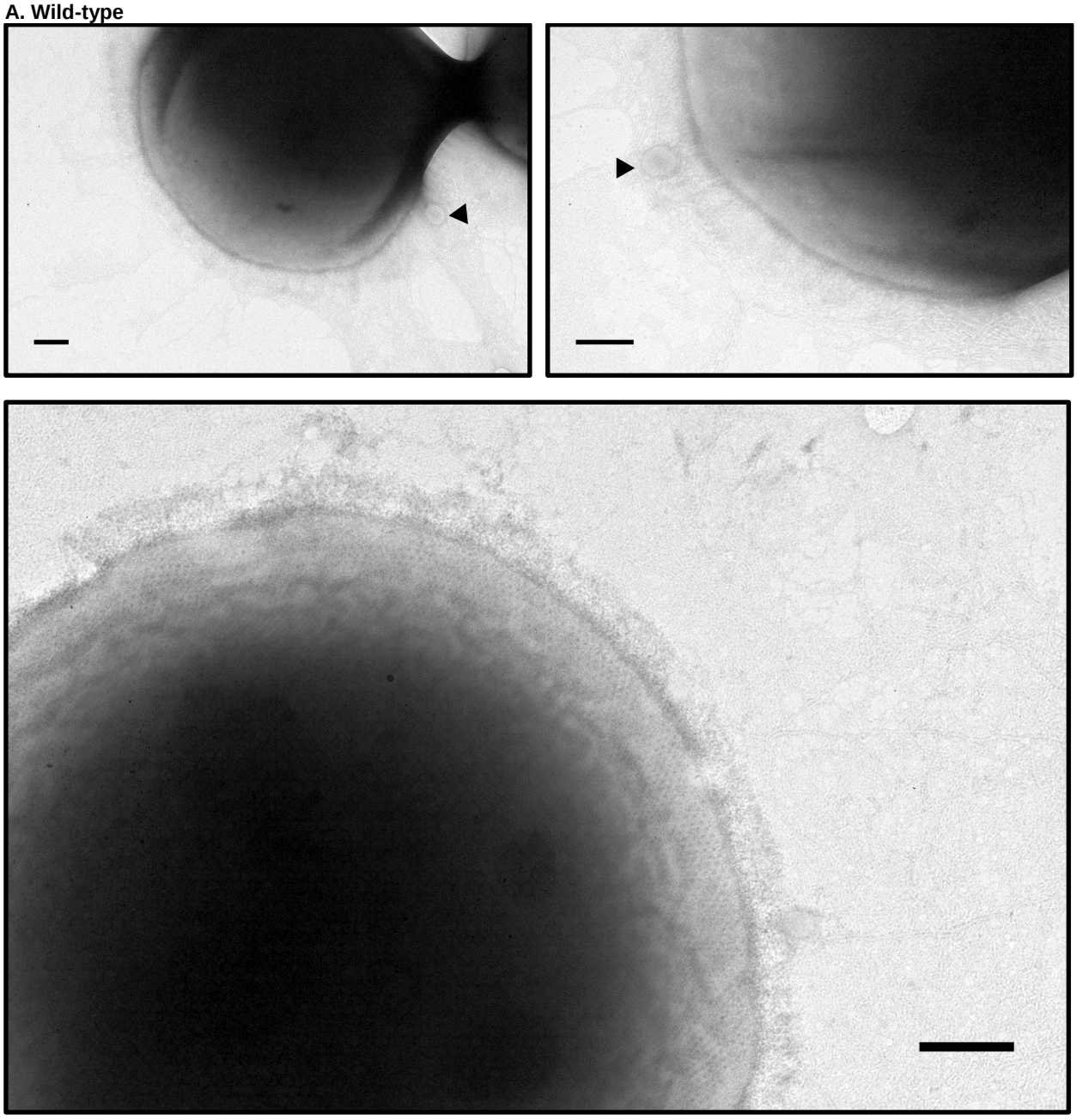
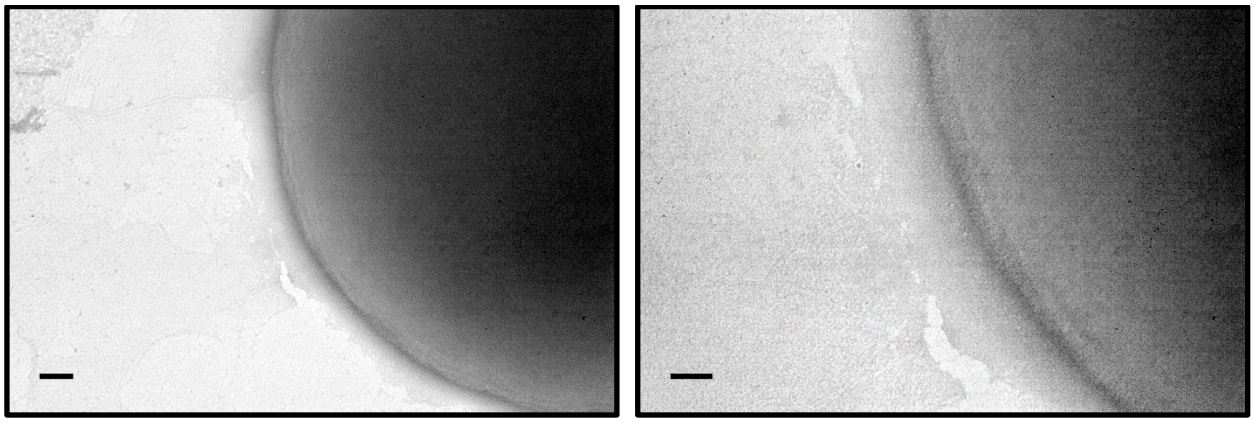


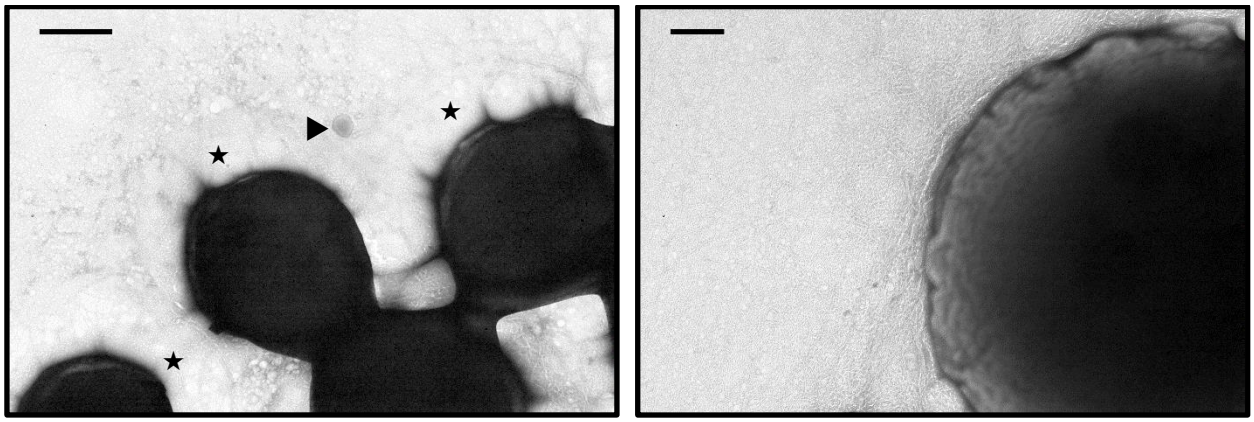
Figure 24: Cell surface analysis of *Synechocystis* wild-type and mutant strains.

Transmission electron micrographs of uranyl acetate negative stained cells of *Synechocystis* wild-type (A), Δ*sll0141* (B), Δ*sll0180* (C), Δ*sll1180* (D), Δ*sll1181* (E), Δ*slr0369* (F), and Δ*slr2131* (G). (A) Wild-type. Upper panel, left: 60000 × amplification, scale bar: 200 nm. Upper panel, right: 100000 × amplification, scale bar: 100 nm. Lower panel: 80000 × amplification, scale bar: 200 nm. It is possible to observe a reticulated lattice as the outermost cell wall structure in wild-type cells, consistent with the formerly described S-layer (*60*, *105*). Vesicles are indicated by arrowheads (▶).

B. Δ*sll0141*



C. Δ*sll0180*



D. Δ*sll1180*

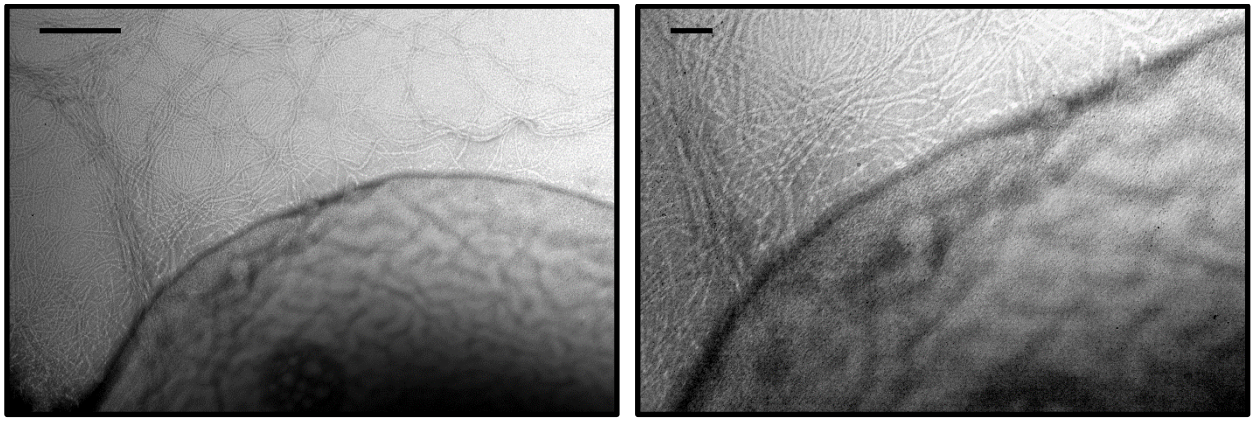
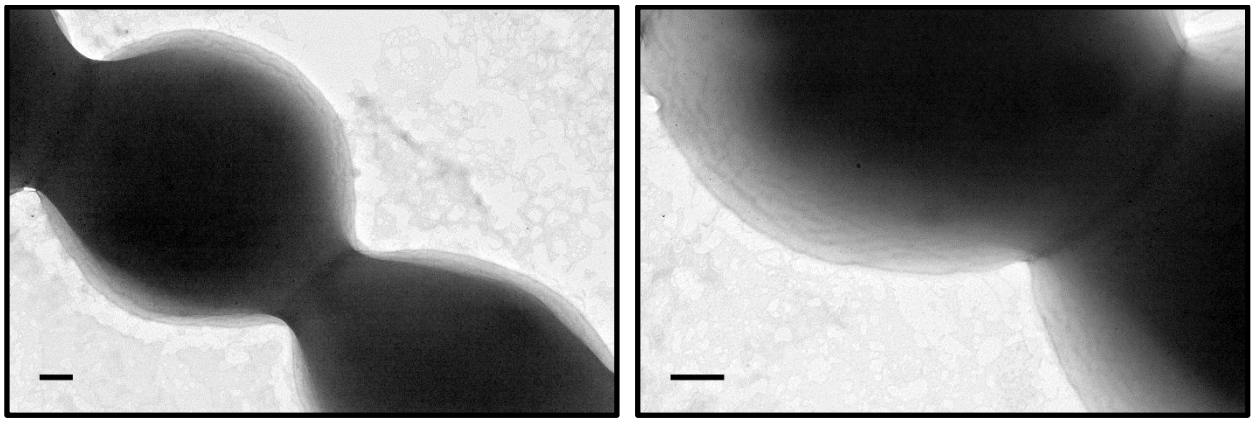


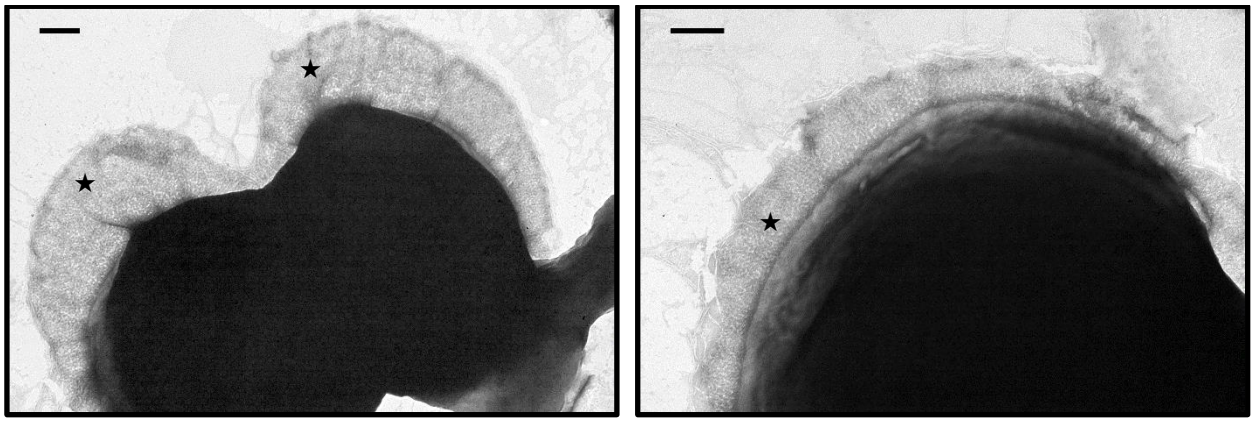
Figure 24: Continued.

(B) Δ*sll0141*. Left: 100000 × amplification, size bar: 100 nm. Right: 250000 × amplification, size bar: 50 nm. (C) Δ*sll0180*. Left: 40000 × amplification, size bar: 550 nm. Right: 80000 × amplification, size bar: 200 nm. (D) Δ*sll1180*. Left: 120000 × amplification, size bar: 200 nm. Right: 250000 × amplification, size bar: 50 nm. Vesicles are indicated by arrowheads (▶). The capsule-like amorphous mass of pili that surrounds the cell is highlighted by a star (★).

E. Δ*sll1181*



F. Δ*slr0369*



G. Δ*slr0454*

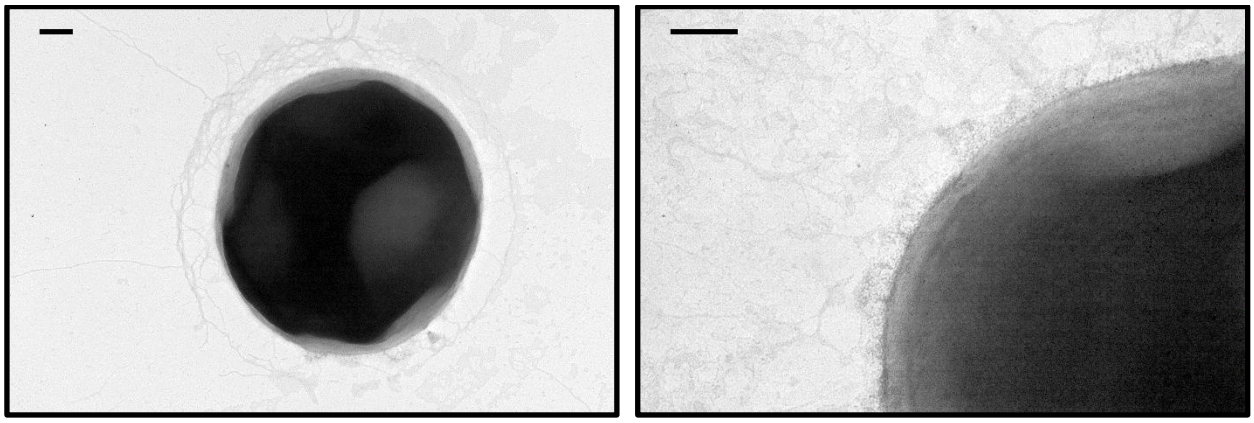


Figure 24: Continued.

(E) Δ*sll1181*. Left: 50000 × amplification, size bar: 200 nm. Right: 80000 × amplification, size bar: 200 nm. (F) Δ*slr0369*. Left: 60000 × amplification, size bar: 200 nm. Right: 80000 × amplification, size bar: 200 nm. (G) Δ*slr0454*. Left: 50000 × amplification, size bar: 200 nm. Right: 100000 × amplification, size bar: 200 nm. Vesicles are indicated by arrowheads. Vesicles are indicated by arrowheads (▶). The capsule-like amorphous mass of pili that surrounds the cell is highlighted by a star (★).

H. Δ*slr2131*

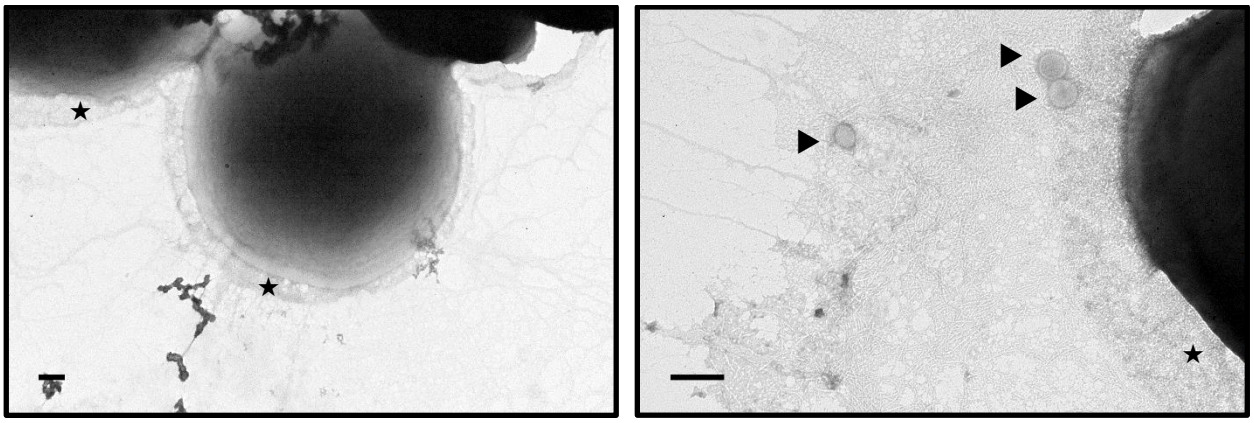


Figure 24: Continued.

(G) Δ*slr2131*. Left: 40000 × amplification, size bar: 200 nm. Right: 80000 × amplification, size bar: 200 nm. Vesicles are indicated by arrowheads (▶). The capsule-like amorphous mass of pili that surrounds the cell is highlighted by a star (★).

4. CONCLUSIONS

4 Conclusions

In the present work, some of the fundamental characteristics of TolC-dependent export mechanisms were studied in the unicellular cyanobacterium *Synechocystis*, and the major conclusions are:

(i) Eleven candidate genes presumptively related to the RND superfamily of multidrug transporters (*sll0141*, *sll0142*, *sll0180*, *sll1053*, *slr0369*, *slr0454*, *slr2131*), the T1SS (*sll1180*, *sll1181*), and homologous to a TolC-dependent glycolipid efflux pump from *Anabaena* sp. PCC 7120 (*sll1481*, *sll1482*) were identified using *in silico* tools;

(ii) *Synechocystis* knockout mutants were constructed by insertional inactivation of the candidate genes, amounting to a total of twelve mutant strains (11 single knockout mutants and 1 double knockout mutant);

(iii) Assessment of the growth of the mutant strains (including Δ*tolC*) under standard laboratory conditions revealed the proteins encoded by the inactivated genes are not essential for cell survival and growth;

(iv) The role of the candidate proteins in the export of different exogenous compounds was investigated by analysing the tolerance of the wild-type and the knockout mutants. In brief, it was demonstrated that mutants lacking functional TolC are hypersensitive towards chloramphenicol, SDS and Triton X-100, and accumulate ethidium bromide. Regarding the remaining knockout mutants, Δ*sll1053*, Δ*slr0454* and Δ*slr2131* presented increased susceptibility to chloramphenicol, Δ*sll0142*, Δ*sll0180* and Δ*slr0369* to SDS and Δ*sll0142*, Δ*sll0180*, Δ*sll1053,* Δ*slr0369* and Δ*slr0454* to Triton X-100, and that Δ*sll1053* and Δ*slr2131* accumulate similar levels of ethidium bromide, as determined for TolC, suggesting the proteins encoded by the disrupted genes are involved in the export of these substances, possibly in a TolC-dependent manner;

(v) The role of the candidate proteins in the secretion of proteins was also studied. Secretion of T1SS substrates depends at least on the function of five proteins: an OM channel-tunnel – TolC, a putative IM transporter encoded by the *sll1180* gene, and three putative PAPs – Sll0141, Sll1080 and Sll1181. Furthermore, in the single knockout mutants Δ*sll0141*, Δ*sll0180*, Δ*slr0369* and Δ*slr2131* glycosylation of the main structural component of pili (PilA1, Sll1694) does not occur, which results in the presence of numerous pili, sometimes forming a capsule-like amorphous mass that encircles the cell;

(vi) The Δ*sll1481*, Δ*sll1482* and Δ*sll1481*Δ*sll1482* knockout mutants did not demonstrate increased susceptibility to any of the exogenous compounds tested, and accumulate ethidium bromide in levels similar to the wild-type. Moreover, the exoproteome of these strains is similar to that of the wild-type. These results suggest that the proteins encoded by the disrupted genes are not involved in chloramphenicol, SDS and Triton X-100 or in protein secretion;

(vii) To sum up, this work provides an important insight into the extensive functionality of TolC-dependent IM translocase complexes in *Synechocystis* and emphasises a TolC-independent functional role between some IM components.

5. FUTURE PERSPECTIVES

5 Future perspectives

In the course of this work, some questions were raised concerning the dynamic interplay of TolC-dependent IM translocase complexes in *Synechocystis*. As described in the literature, promiscuous partnering and substrate recognition are typical features of these systems, adding additional layers of functional and mechanistic complexity to these multiprotein assemblies, which have often been characterised as the simplest export system capable of traversing the Gram-negative cell envelope. Experimental evidences provided in this work, point towards the participation of five proteins in the T1SS (TolC, Slr1270; an IMC, Sll1181; and three PAPs, Sll0141, Sll0180, Sll1181). Moreover, it has been demonstrated that Sll0141 and Sll0180 also partake in other mechanisms, namely the glycosylation of the main structural component of pili, the protein PilA1. In this context, *in vivo* cross-linking experiments would help disclose the network of protein-protein interactions responsible for the dynamic interplay of proteins in TolC-dependent export mechanisms. The information obtained from cross-linking experiments would provide essential data about the structure and function of the herein studied proteins to the point it allows mapping of protein-protein interactions to specific amino acids or domains.

Synechocystis, one of the best studied cyanobacteria, is a strong candidate for the photosynthetic and sustainable production of high-value chemicals, namely short chain alcohols (e.g. ethanol, 2-propanol and 1-butanol). It has been acknowledged that bacterial efflux systems are crucial to improve the overall production yield, and TolC-dependent membrane transporters have been successfully engineered to improve tolerance and productivity in other Gram-negative bacteria. For this reason, a screening for transporters capable of secreting added-value compounds would help develop a cyanobacterial strain better equipped to withstand the detrimental effects exerted by the compound, increasing production titers. In the future, the knockout mutants constructed in this work could help identify which transporters are capable of secreting compounds of biotechnological interest. Preliminary results indicate that the Δ*tolC* knockout mutant is approximately 16 times more tolerant than the wild-type, when exposed to 0.125 % (v/v) 1-butanol for 7 days (Supplementary data, **Figure S1**). Inactivation of TolC impairs the cellular capacity for detoxification and repair; nonetheless, the Δ*tolC* knockout mutant proved to be more resistant than the wild-type to an exogenous compound whose biochemical nature renders it toxic to cells. A thorough understanding of the cellular adaptations resulting from the absence of TolC and exposure to 1-butanol must include an analysis of lipid composition, as it has been shown in other microorganisms that the membrane' structural components change

dramatically when exposed to alcohols. Moreover, it has been demonstrated that mutants lacking functional TolC suffer oxidative damage to membranes and enzymatic systems responsible for the detoxification of reactive oxygen species present higher activity levels. Therefore, whole-cell lipidomics studies would provide vital information about the lipid composition adaptations occurring in cyanobacterial membranes upon exposure to alcohols, which constitutes an important step for the optimisation of cyanobacteria as cell factories.

6. REFERENCES

6 References

1. B. Rasmussen, I. R. Fletcher, J. J. Brocks, M. R. Kilburn, Reassessing the first appearance of eukaryotes and cyanobacteria. *Nature*. **455**, 1101–1104 (2008).

2. K. Sciuto, I. Moro, Cyanobacteria: the bright and dark sides of a charming group. *Biodivers. Conserv.* **24**, 711–738 (2015).

3. A. M. Ruffing, Engineered cyanobacteria: teaching an old bug new tricks. *Bioeng. Bugs*. **2**, 136–149 (2011).

4. B. E. Schirrmeister, A. Antonelli, H. C. Bagheri, The origin of multicellularity in cyanobacteria. *BMC Evol. Biol.* **11**, 45–65 (2011).

5. E. Dittmann *et al.*, Natural product biosynthetic diversity and comparative genomics of the cyanobacteria. *Trends Microbiol.* **23**, 642–652 (2015).

6. S. V. Shestakov, N. T. Khyen, Evidence for genetic transformation in blue-green alga *Anacystis nidulans*. *MGG Mol. Gen. Genet.* **107**, 372–375 (1970).

7. G. Grigorieva, S. Shestakov, Transformation in the cyanobacterium *Synechocystis* sp. 6803. *FEMS Microbiol. Lett.* **13**, 367–370 (1982).

8. Z. Gao, H. Zhao, Z. Li, X. Tan, X. Lu, Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria. *Energy Environ. Sci.* **5**, 9857–9865 (2012).

9. C. R. Shen, J. C. Liao, Photosynthetic production of 2-methyl-1-butanol from CO_2 in cyanobacterium *Synechococcus elongatus* PCC7942 and characterization of the native acetohydroxyacid synthase. *Energy Environ. Sci.* **5**, 9574 (2012).

10. Y. Xue, Q. He, Cyanobacteria as cell factories to produce plant secondary metabolites. *Front. Bioeng. Biotechnol.* **3**, 57 (2015).

11. S. A. Angermayr, A. Gorchs Rovira, K. J. Hellingwerf, Metabolic engineering of cyanobacteria for the synthesis of commodity products. *Trends Biotechnol.* **33**, 352–361 (2015).

12. J. Zhou, T. Zhu, Z. Cai, Y. Li, From cyanochemicals to cyanofactories: a review and perspective. *Microb. Cell Fact.* **15**, 2 (2016).

13. C. T. Trinh, P. Unrean, F. Srienc, Minimal *Escherichia coli* cell for the most efficient

production of ethanol from hexoses and pentoses. *Appl. Environ. Microbiol.* **74**, 3634–3643 (2008).

14. N. S. Lau, M. Matsui, A. A. A. Abdullah, Cyanobacteria: photoautotrophic microbial factories for the sustainable synthesis of industrial products. *Biomed Res. Int.* **2015** (2015), doi:10.1155/2015/754934.

15. K. Kanekiyo *et al.*, Isolation of an antiviral polysaccharide, Nostoflan, from a terrestrial cyanobacterium, Nostoc flagelliforme. *J. Nat. Prod.* **68**, 1037–1041 (2005).

16. J. W. J. J. W. J. Lee *et al.*, Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat. Chem. Biol.* **8**, 536–546 (2012).

17. V. Chubukov, A. Mukhopadhyay, C. J. Petzold, J. D. Keasling, H. G. Martín, Synthetic and systems biology for microbial production of commodity chemicals. *Syst. Biol. Appl.* **2**, 16009 (2016).

18. D. Kell, Metabolic control theory: its role in microbiology and biotechnology. *FEMS Microbiol. Lett.* **39**, 305–320 (1986).

19. D. B. Kell, N. Swainston, P. Pir, S. G. Oliver, Membrane transporter engineering in industrial biotechnology and whole cell biocatalysis. *Trends Biotechnol.* **33**, 237–246 (2015).

20. T. K. Van Dyk, in *Advances in Applied Microbiology* (2008), vol. 63, pp. 231–247.

21. A. Rojas *et al.*, Three efflux pumps are required to provide efficient tolerance to toluene in *Pseudomonas putida* DOT-T1E. *J. Bacteriol.* **183**, 3967–3973 (2001).

22. R. M. Lennen, M. G. Politz, M. A. Kruziki, B. F. Pfleger, Identification of transport proteins involved in free fatty acid efflux in *Escherichia coli*. *J. Bacteriol.* **195**, 135–144 (2013).

23. M. A. Fisher *et al.*, Enhancing tolerance to short-chain alcohols by engineering the *Escherichia coli* AcrB efflux pump to secrete the non-native substrate n-butanol. *ACS Synth. Biol.* **3**, 30–40 (2014).

24. J. Foo, S. Leong, Directed evolution of an *E. coli* inner membrane transporter for improved efflux of biofuel molecules. *Biotechnol. Biofuels*. **6**, 81 (2013).

25. M. R. J. J. Salton, The bacterial cell envelope - a historical perspective. *Cell*. **27**, 1–22 (1994).

26. G. M. Cooper, R. E. Hausman, in *The cell: a molecular approach* (Sinauer Associates, ed.

4th, 2007), pp. 3–15.

27. T. J. Silhavy, D. Kahne, S. Walker, The bacterial cell envelope. *Cold Spring Harb. Perspect. Biol.* **2**, a000414–a000414 (2010).

28. E. Hoiczyk, A. Hansel, Cyanobacterial cell walls: news from an unusual prokaryotic envelope. *J. Bacteriol.* **182**, 1191–1199 (2000).

29. E. Hoiczyk, W. Baumeister, Envelope structure of four gliding filamentous cyanobacteria. *J. Bacteriol.* **177**, 2387–2395 (1995).

30. M. H. Saier, A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol. Mol. Biol. Rev.* **64**, 354–411 (2000).

31. C. Andersen, Channel-tunnels: outer membrane components of type I secretion systems and multidrug efflux pumps of Gram-negative bacteria. *Rev. Physiol. Biochem. Pharmacol*. **147**, 122–165 (2003).

32. V. Koronakis, J. Eswaran, C. Hughes, Structure and function of TolC: the bacterial exit duct for proteins and drugs. *Annu. Rev. Biochem.* **73**, 467–489 (2004).

33. J. M. Johnson, G. M. Church, Alignment and structure prediction of divergent protein families: periplasmic and outer membrane proteins of bacterial efflux pumps. *J. Mol. Biol.* **287**, 695–715 (1999).

34. T. R. D. Costa *et al.*, Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat. Rev. Microbiol.* **13**, 343–359 (2015).

35. J. A. Delmar, C. C. Su, E. W. Yu, Bacterial multidrug efflux transporters. *Annu Rev Biophys*. **43**, 93–117 (2014).

36. J. Anes, M. P. McCusker, S. Fanning, M. Martins, The ins and outs of RND efflux pumps in Escherichia coli. *Front. Microbiol.* **6**, 587 (2015).

37. H. I. Zgurskaya, G. Krishnamoorthy, A. Ntreh, S. Lu, Mechanism and function of the outer membrane channel TolC in multidrug resistance and physiology of enterobacteria. *Front. Microbiol.* **2**, 1–13 (2011).

38. M. F. Symmons, R. L. Marshall, V. N. Bavro, Architecture and roles of periplasmic adaptor proteins in tripartite efflux assemblies. *Front. Microbiol.* **6**, 1–20 (2015).

39. V. Koronakis, A. Sharff, E. Koronakis, B. Luisi, C. Hughes, Crystal structure of the bacterial

membrane protein TolC central to multidrug efflux and protein export. *Nature*. **405**, 914–9 (2000).

40. L. J. V Piddock, Multidrug-resistance efflux pumps - not just for resistance. *Nat. Rev. Microbiol.* **4**, 629–636 (2006).

41. P. Blanco *et al.*, Bacterial multidrug efflux pumps: much more than antibiotic resistance determinants. *Microorganisms*. **4**, 14 (2016).

42. H. I. Zgurskaya, Y. Yamada, E. B. Tikhonova, Q. Ge, G. Krishnamoorthy, Structural and functional diversity of bacterial membrane fusion proteins. *Biochim. Biophys. Acta - Proteins Proteomics*. **1794**, 794–807 (2009).

43. A. T. Rêgo, V. Chandran, G. Waksman, Two-step and one-step secretion mechanisms in Gram-negative bacteria: contrasting the type IV secretion system and the chaperone-usher pathway of pilus biogenesis. *Biochem. J.* **425**, 475–488 (2010).

44. K. V. Korotkov, M. Sandkvist, W. G. J. Hol, The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat. Rev. Microbiol.* (2012), doi:10.1038/nrmicro2762.

45. F. Amaral, C. Jácome, M. Ângela, A. Tillmann, F. Ufpel, Bacterial Secretion Systems – An overview. **4**, 1–32 (2015).

46. J.-S. Kim *et al.*, Crystal structure of a soluble fragment of the membrane fusion protein HlyD in a type I secretion system of Gram-Negative bacteria. *Structure*. **24**, 477–485 (2016).

47. S. Moslavac *et al.*, A TolC-like protein is required for heterocyst development in anabaena sp. strain PCC 7120. *J. Bacteriol.* **189**, 7887–7895 (2007).

48. P. Staron, K. Forchhammer, I. Maldener, Novel ATP-driven pathway of glycolipid export involving TolC protein. *J. Biol. Chem.* **286**, 38202–38210 (2011).

49. P. Staron, K. Forchhammer, I. Maldener, Structure-function analysis of the ATP-driven glycolipid efflux pump DevBCA reveals complex organization with TolC/HgdD. *FEBS Lett.* **588**, 395–400 (2014).

50. A. Hahn, M. Stevanovic, O. Mirus, E. Schleiff, The TolC-like protein HgdD of the cyanobacterium *Anabaena* sp. PCC 7120 is involved in secondary metabolite export and antibiotic resistance. *J. Biol. Chem.* **287**, 41126–41138 (2012).

51. A. Hahn *et al.*, Secretome analysis of *Anabaena* sp: PCC 7120 and the involvement of the TolC-homologue HgdD in protein secretion. *Environ. Microbiol.* **17**, 767–780 (2015).

52. K. Nicolaisen *et al.*, The interplay between siderophore secretion and coupled iron and copper transport in the heterocyst-forming cyanobacterium *Anabaena* sp. PCC 7120. *Biochim. Biophys. Acta - Biomembr.* **1798**, 2131–2140 (2010).

53. M. Stevanovic, A. Hahn, K. Nicolaisen, O. Mirus, E. Schleiff, The components of the putative iron transport system in the cyanobacterium *Anabaena* sp. PCC 7120. *Environ. Microbiol.* **14**, 1655–1670 (2012).

54. A. Kato *et al.*, Identification of a cyanobacterial RND-type efflux system involved in export of free fatty acids. *Plant Cell Physiol.* **56**, 2467–2477 (2015).

55. W. Vermaas, Molecular genetics of the cyanobacterium *Synechocystis* sp. PCC 6803: principles and possible biotechnology applications. *J. Appl. Phycol.* **8**, 263–273 (1996).

56. T. Kaneko *et al.*, Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* **3**, 109–136 (1996).

57. T. Kaneko, Structural analysis of four large plasmids harboring in a unicellular cyanobacterium, *Synechocystis* sp. PCC 6803. *DNA Res.* **10**, 221–228 (2003).

58. Y. Nakamura, CyanoBase, the genome database for *Synechocystis* sp. strain PCC6803: status for the year 2000. *Nucleic Acids Res.* **28**, 72–72 (2000).

59. R. Agarwal *et al.*, Structure-function of cyanobacterial outer-membrane protein, Slr1270: homolog of *Escherichia coli* drug export/colicin import protein, TolC. *FEBS Lett.* **588**, 3793–3801 (2014).

60. P. Oliveira *et al.*, The versatile TolC-like Slr1270 in the cyanobacterium *Synechocystis* sp. PCC 6803. *Environ. Microbiol.* (2015), doi:10.1111/1462-2920.13172.

61. R. Y. Stanier, R. Kunisawa, M. Mandel, G. Cohen-Bazire, Purification and properties of unicellular blue-green algae (order Chroococcales). *Bacteriol. Rev.* **35**, 171–205 (1971).

62. G. Bertani, Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* **62**, 293–300 (1951).

63. J. C. Meeks, R. W. Castenholz, Growth and photosynthesis in an extreme thermophile,

Synechococcus lividus (Cyanophyta). *Arch. Mikrobiol.* **78**, 25–41 (1971).

64. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–10 (1990).

65. A. Marchler-Bauer *et al.*, CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **43**, D222–D226 (2015).

66. L. Käll, A. Krogh, E. L. . Sonnhammer, A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* **338**, 1027–1036 (2004).

67. A. S. Juncker *et al.*, Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci.* **12**, 1652–1662 (2003).

68. T. Koressaar, M. Remm, Enhancements and modifications of primer design program Primer3. *Bioinformatics*. **23**, 1289–1291 (2007).

69. P. Tamagnini, O. Troshina, F. Oxelfelt, R. Salema, P. Lindblad, Hydrogenases in *Nostoc* sp. strain PCC 73102, a strain lacking a bidirectional enzyme. *Appl. Envir. Microbiol.* **63**, 1801–1807 (1997).

70. D. Hanahan, J. Jessee, F. R. Bloom, *Bacterial Genetic Systems* (Elsevier, 1991), vol. 204 of *Methods in Enzymology*.

71. D. Hanahan, Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**, 557–580 (1983).

72. T. Heidorn *et al.*, in *Methods in Enzymology* (2011), vol. 497, pp. 539–579.

73. P. Oliveira, F. Pinto, C. C. Pacheco, R. Mota, P. Tamagnini, HesF, an exoprotein required for filament adhesion and aggregation in *Anabaena* sp. PCC 7120. *Environ. Microbiol.* **17**, 1631–1648 (2015).

74. H. Okusu, D. Ma, H. Nikaido, AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of *Escherichia coli* multiple-antibiotic-resistance (Mar) mutants. *J. Bacteriol.* **178**, 306–308 (1996).

75. T. Touzé *et al.*, Interactions underlying assembly of the *Escherichia coli* AcrAB-TolC multidrug efflux system. *Mol. Microbiol.* **53**, 697–706 (2004).

76. D. Du *et al.*, Structure of the AcrAB–TolC multidrug efflux pump. *Nature*. **509**, 512–515

(2014).

77. C. Ruiz, S. B. Levy, Regulation of acrAB expression by cellular metabolites in *Escherichia coli*. *J. Antimicrob. Chemother.* **69** (2014), doi:10.1093/jac/dkt352.

78. A. S. Jellen-Ritter, W. V. Kern, Enhanced expression of the multidrug efflux pumps AcrAB and AcrEF associated with insertion element transposition in *Escherichia coli* mutants selected with a fluoroquinolone. *Antimicrob. Agents Chemother.* **45**, 1467–1472 (2001).

79. S. Y. Lau, H. I. Zgurskaya, Cell division defects in *Escherichia coli* deficient in the multidrug efflux transporter AcrEF-TolC. *J. Bacteriol.* **187**, 7815–7825 (2005).

80. C. Alvarez-Ortega, J. Olivares, J. L. Martínez, RND multidrug efflux pumps: what are they good for? *Front. Microbiol.* **4** (2013), p. 7.

81. D. C. Rees, E. Johnson, O. Lewinson, ABC transporters: the power to change. *Nat. Rev. Mol. Cell Biol.* **10**, 218–227 (2009).

82. C. K. W. Schwarz *et al.*, Crystallization and preliminary X-ray crystallographic studies of an oligomeric species of a refolded C39 peptidase-like domain of the *Escherichia coli* ABC transporter haemolysin B. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **67**, 630–633 (2011).

83. T. Yakushi, K. Masuda, S. Narita, S. Matsuyama, H. Tokuda, A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. *Nat. Cell Biol.* **2**, 212–8 (2000).

84. S. K. Buchanan, Type I secretion and multidrug efflux: transport through the TolC channel-tunnel. *Trends Biochem. Sci.* **26**, 3–6 (2001).

85. I. B. Holland, L. Schmitt, J. Young, Type 1 protein secretion in bacteria, the ABC-transporter dependent pathway (review). *Mol. Membr. Biol.* **22**, 29–39 (2005).

86. L. Schmitt, H. Benabdelhak, M. A. Blight, I. B. Holland, M. T. Stubbs, Crystal structure of the nucleotide-binding domain of the ABC-transporter haemolysin B: identification of a variable region within ABC helical domains. *J. Mol. Biol.* **330**, 333–342 (2003).

87. I. B. Holland, M. A. Blight, B. Kenny, The mechanism of secretion of hemolysin and other polypeptides from Gram-negative bacteria. *J. Bioenerg. Biomembr.* **22**, 473–91 (1990).

88. I. T. Paulsen, M. H. Brown, R. A. Skurray, Proton-dependent multidrug efflux systems.

Microbiol. Rev. **60**, 575–608 (1996).

89. O. Lomovskaya, H. I. Zgurskaya, H. Nikaido, It takes three to tango. *Nat. Biotechnol.* **20**, 1210–1212 (2002).

90. O. Lomovskaya, K. Lewis, Emr, an *Escherichia coli* locus for multidrug resistance. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 8938–42 (1992).

91. L. Balakrishnan, C. Hughes, V. Koronakis, Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by *Escherichia coli*. *J. Mol. Biol.* **313**, 501–10 (2001).

92. W. Mao *et al.*, On the mechanism of substrate specificity by resistance nodulation division (RND)-type multidrug resistance pumps: the large periplasmic loops of MexD from *Pseudomonas aeruginosa* are involved in substrate recognition. *Mol. Microbiol.* **46**, 889–901 (2002).

93. C. A. Elkins, H. Nikaido, Substrate specificity of the RND-type multidrug efflux pumps AcrB and AcrD of *Escherichia coli* is determined predominately by two large periplasmic loops. *J. Bacteriol.* **184**, 6490–6498 (2002).

94. J. Labarre, F. Chauvat, P. Thuriaux, Insertional mutagenesis by random cloning of antibiotic resistance genes into the genome of the cyanobacterium *Synechocystis* strain PCC 6803. *J. Bacteriol.* **171**, 3449 (1989).

95. K. Poole, Efflux-mediated multiresistance in Gram-negative bacteria. *Clin. Microbiol. Infect.* **10**, 12–26 (2004).

96. C. Wang, L. Yang, A. A. Shah, E.-S. Choi, S.-W. Kim, Dynamic interplay of multidrug transporters with TolC for isoprenol tolerance in *Escherichia coli*. *Sci. Rep.* **5**, 16505 (2015).

97. J. F. Linares, I. Gustafsson, F. Baquero, J. L. Martinez, Antibiotics as intermicrobial signaling agents instead of weapons. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 19484–9 (2006).

98. G. Yim, H. Huimi Wang, J. Davies FRS, Antibiotics as signalling molecules. *Philos. Trans. R. Soc. London B Biol. Sci.* **362** (2007).

99. J. Davies, G. B. Spiegelman, G. Yim, The world of subinhibitory antibiotic concentrations. *Curr. Opin. Microbiol.* **9**, 445–453 (2006).

100. L. Migliore, A. Rotini, M. C. Thaller, Low doses of tetracycline trigger the *E. coli* growth: a

case of hormetic response. *Dose. Response.* **11**, 550–7 (2013).

101. J. D. Desai, I. M. Banat, Microbial production of surfactants and their commercial potential. *Microbiol. Mol. Biol. Rev.* **61**, 47–64 (1997).

102. E. B. Tikhonova, V. Dastidar, V. V Rybenkov, H. I. Zgurskaya, Kinetic control of TolC recruitment by multidrug efflux complexes. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 16416–21 (2009).

103. I. Couto, S. S. Costa, M. Viveiros, M. Martins, L. Amaral, Efflux-mediated response of *Staphylococcus aureus* exposed to ethidium bromide. *J. Antimicrob. Chemother.* **62**, 504–513 (2008).

104. M. Ferhat *et al.*, The TolC Protein of *Legionella pneumophila* plays a major role in multi-drug resistance and the early steps of host invasion. *PLoS One*. **4**, e7732 (2009).

105. C. Trautner, W. F. J. Vermaas, The sll1951 gene encodes the surface layer protein of *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.* **195**, 5370–5380 (2013).

106. G. Shen, S. Boussiba, W. F. Vermaas, *Synechocystis* sp PCC 6803 strains lacking photosystem I and phycobilisome function. *Plant Cell*. **5**, 1853–63 (1993).

107. J. M. Engle, W. Burkhart, D. M. Sherman, G. S. Bullerjahn, Purification and characterization of a surface-associated carotenoid-binding complex from the photosynthetic prokaryote, *Prochlorothrix hollandica*. *Arch. Microbiol.* **155**, 453–458 (1991).

108. T. Mima, S. Joshi, M. Gomez-Escalada, H. P. Schweizer, Identification and characterization of TriABC-OpmH, a triclosan efflux pump of *Pseudomonas aeruginosa* requiring two membrane fusion proteins. *J. Bacteriol.* **189**, 7600–7609 (2007).

109. J. W. Weeks, L. M. Nickels, A. T. Ntreh, H. I. Zgurskaya, Non-equivalent roles of two periplasmic subunits in the function and assembly of triclosan pump TriABC from *Pseudomonas aeruginosa*. *Mol. Microbiol.* **98**, 343–356 (2015).

110. T. Gristwood, P. C. Fineran, L. Everson, G. P. C. Salmond, PigZ, a TetR/AcrR family repressor, modulates secondary metabolism via the expression of a putative four-component resistance-nodulation-cell-division efflux pump, ZrpADBC, in *Serratia* sp. ATCC 39006. *Mol. Microbiol.* **69**, 418–435 (2008).

111. J. Mikolosko, K. Bobyk, H. I. Zgurskaya, P. Ghosh, Conformational flexibility in the multidrug

efflux system protein AcrA. *Structure*. **14**, 577–587 (2006).

112. N. Schuergers, A. Wilde, Appendages of the cyanobacterial cell. *Life (Basel, Switzerland)*. **5**, 700–15 (2015).

113. Y. H. Kim *et al.*, The role of Slr1443 in pilus biogenesis in *Synechocystis* sp. PCC 6803: involvement in post-translational modification of pilins. *Biochem. Biophys. Res. Commun.* **315**, 179–186 (2004).

114. Y. H. Kim *et al.*, Alteration in the glycan pattern of pilin in a nonmotile mutant of *Synechocystis* sp. PCC 6803. *Proteomics*. **9**, 1075–86 (2009).

115. Y. H. Kim *et al.*, Identification of trimethylation at C-terminal lysine of pilin in the cyanobacterium *Synechocystis* PCC 6803. *Biochem. Biophys. Res. Commun.* **404**, 587–92 (2011).

7. SUPPLEMENTARY DATA

7 Supplementary data

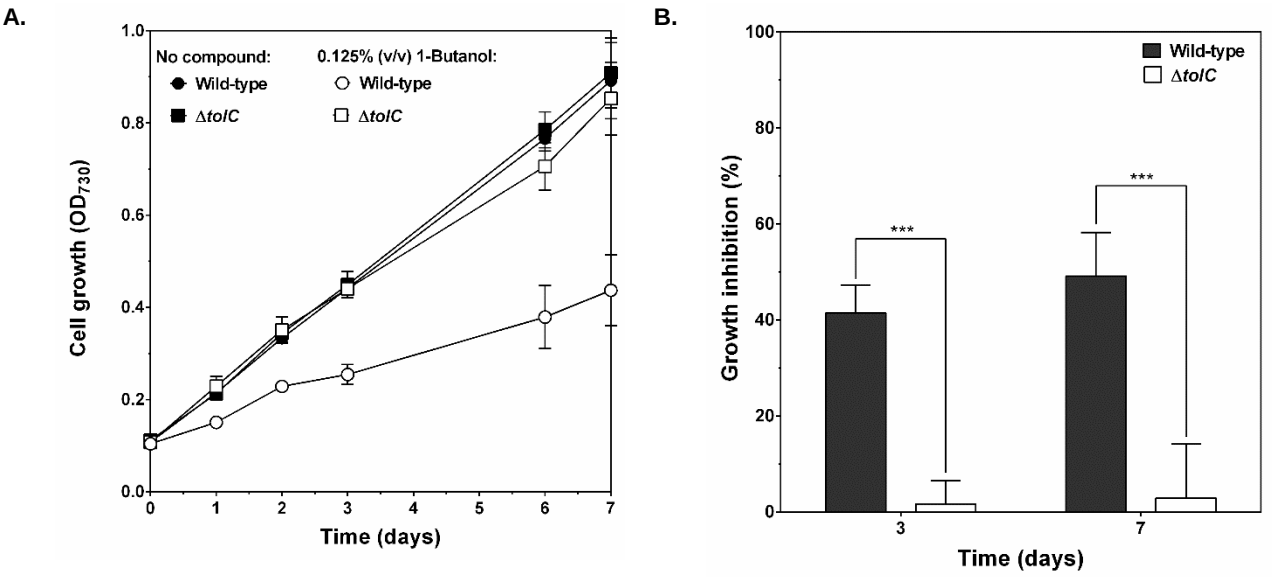


Figure S1: Assessment of the growth fitness of the *Synechocystis* wild-type and *ΔtolC* mutant when exposed to 1-butanol.

Synechocystis wild-type and mutant strains were grown for 7 days in 500 mL Erlenmeyer flasks, at 25 °C, under a regimen of 16h light (25 µmol photons m^{-2} s^{-1}))/8h dark regimen, in BG11 medium or BG11 medium supplemented 0.125 % (v/v) 1-butanol. (A) Cell growth (OD_{730}) was measured for a period of 7 days. (B) The growth inhibition percentage was calculated as a ratio between the growth rate of the strain in the presence of the compound and in its absence. Error bars represent the standard deviations of three independent biological replicates. ns, P > 0.05; *, P ≤ 0.05; **, P ≤ 0.01; ***, P ≤ 0.001; ****, P ≤ 0.0001.